Structural studies of two enzymes in the Raetz pathway of lipid A synthesis, LpxB and LpxH

A DISSERTATION SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF UNIVERSITY OF MINNESOTA

BY

Thomas E. Bohl

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

ADVISER: Hideki Aihara

May, 2018

Acknowledgements

I would like to thank the people in the Aihara Lab who made my research possible. In particular, my thanks goes to my advisor Dr. Hideki Aihara who has guided me through the research and publication processes. Hideki has always been available to discuss how to address problems I faced during my research and how to communicate the results of my research effectively in my publications. At the same time, Hideki has always given me the freedom to take the project where I wanted to. I also want to thank Dr. John Lee who got me started researching these Raetz pathway enzymes and who has helped me appreciate the place of my research in the history of research on lipid A synthesis. In addition, I thank Dr. Ke Shi for his invaluable help with x-ray diffraction, and I thank Kayo Kurahashi for her help in cloning.

I would also like to thank my collaborators in Dr. Rommie Amaro's lab at University of California: San Diego, Pek Ieong and Dr. Özlem Demir. Not only did they perform the molecular dynamics simulations for the LpxH project, but our discussions with them helped shape our understanding of the role of flexibility in the LpxH cap.

Other collaborators were also instrumental in this research. I thank Dr. Jayakanth Kankanala of Dr. Zhengqiang Wang's lab in the Center for Drug Design for his help with organic synthesis and especially for his availability for help with mass spectrometry. In addition, I thank Dr. Surajit Banerjee from the Advanced Photon Source at Argonne National Lab for his excellent assistance with x-ray diffraction. I would also like to thank Dr. Kurt Peterson of Fluorescence Innovations whose in-house built luminometer made the LpxH activity assays possible, and I would like to thank Dr. Thomas Lee of

University of Colorado: Boulder who performed and analyzed hydrogen-deuterium exchange mass spectrometry.

Finally, I would like to thank my classmates who have made the graduate experience fun. I couldn't have attended graduate school with a better group of people.

Dedication

I dedicate this dissertation to Boris and Natasha: my companions, my confidants, my alarm clock, my beloved feline overlords.

Dissertation Abstract

Gram-negative bacteria are distinguished from Gram-positive bacteria by the secondary membrane that surrounds their peptidoglycan cell wall. The outer leaflet of this membrane is primarily composed of the glycolipid lipopolysaccharide (LPS), which has lipid A, core oligosaccharide, and O-antigen portions. LPS helps protect Gram-negative bacteria from hydrophobic toxins and, in pathogenic bacteria, from the host immune system. The membrane anchor portion of LPS (lipid A) is responsible for stimulation of the inflammatory response of the mammalian immune system by LPS via activation of the Toll-like receptor 4/myeloid differentiation factor 2 complex. In systemic infections, overstimulation of this receptor causes acute inflammation, which can cause septic shock. Modifications to the LPS, particularly to the lipid A portion, can help bacteria evade the host immune system by disguising the bacteria, modulating the inflammatory response, and inhibiting interactions with antimicrobial host factors.

Lipid A is synthesized in the well characterized and largely conserved Raetz pathway in the cytosol and at the cytosolic face of the inner membrane. The non-repeating core oligosaccharide is synthesized on lipid A at the cytoplasmic face of the inner membrane, and the repeating O-antigen polysaccharide is attached to the core-lipid A molecule at the periplasmic face of the inner membrane. The completed LPS molecules are then transported to the extracellular leaflet of the outer membrane. Structures of proteins involved in LPS synthesis have proved critical to our understanding of LPS synthesis and transport. Moreover, these structures provide targets for rational design of antibiotics targeting Gram-negative bacteria. As the Raetz pathway is the most conserved part of LPS synthesis, the enzymes of the Raetz pathway provide particularly promising

targets for development of broad spectrum antibiotics, such as those needed to treat sepsis. Therefore, I studied the structures of two enzymes in the Raetz pathway (LpxH and LpxB).

In the crystal structure of *E. coli* LpxH, crystal packing captured LpxH with the α-helical substrate-binding cap domain in a displaced conformation, suggesting that this domain is highly mobile. The structural dynamics of this domain and their relevance to substrate binding were further explored by hydrogen-deuterium exchange mass spectrometry, molecular dynamics simulations, and activity assays. These data supported a model in which a loop in the core hydrolase domain acts as a wedge to promote an opening motion of the capping helices wherein bending and partial unwinding of the helices exposes the active site. This opening may allow facile substrate binding between these helices.

The first structure of LpxB was determined showing a Glycosyltransferase B superfamily (GT-B) fold modified by the formation of a novel C-terminally swapped dimer wherein the last 87 residues of one subunit complete the GT-B fold of the other subunit. Furthermore, the binding site of the sugar-donor substrate was identified by a structure of LpxB with the UDP product bound. Activity assays supported the formation of this C-terminally swapped dimer in solution and showed that a surface-exposed hydrophobic patch is critical for LpxB activity, which suggested this patch allows productive membrane association required for substrate binding. Thus, the present research has expanded our understanding of two enzymes important in Gram-negative bacterial physiology. These enzymes are potential targets for antibiotic development.

Table of Contents

List of Tables

List of Figures

List of Schemes

CHAPTER 1: INTRODUCTION

1.1 Background and Significance

Gram-negative bacteria are distinguished from Gram-positive bacteria by the secondary lipid bilayer that surrounds their peptidoglycan cell wall (1). In the majority of Gram-negative bacteria, this outer membrane is an asymmetric bilayer with phospholipids on the inner leaflet and lipopolysaccharide (LPS) on the outer leaflet (1, 2). LPS is a glycolipid composed of a lipid A membrane anchor, a core oligosaccharide, and a repeating O antigen polysaccharide (Figure 1.1) (1, 3). This layer of LPS provides a permeability barrier to environmental toxins. When this layer of LPS is disrupted by invasion of phospholipids (2), loss of LPS synthesis or transport (4), or loss of glycosylation (5), the bacteria show increased sensitivity to hydrophobic toxins, such as detergent or bile salts, or antibiotics. Furthermore, complete loss of LPS by mutation of lipid A synthesis genes made bacteria less virulent to *Caenorhabditis elegans* and human epithelial cells (6). Likely in part because of these deleterious effects, loss of lipid A synthesis has only been characterized in *Acinetobacter baumannii* (4, 6), and four of the five enzymes responsible for synthesis of the base tetra-acylated disaccharide are conserved in all but the specialized Gram-negative bacteria, such as *Sphingomonads*, that use alternative lipids (7, 8).

The role of LPS in disease is related not only to its protective function in Gram-negative bacteria but also to its interaction with the host's innate immune system. Humans possess a cohort of proteins that allow rapid response to the presence of LPS (9). LPS binding protein is thought to extract LPS from bacterial membranes and present an LPS molecule to CD14 (9, 10).

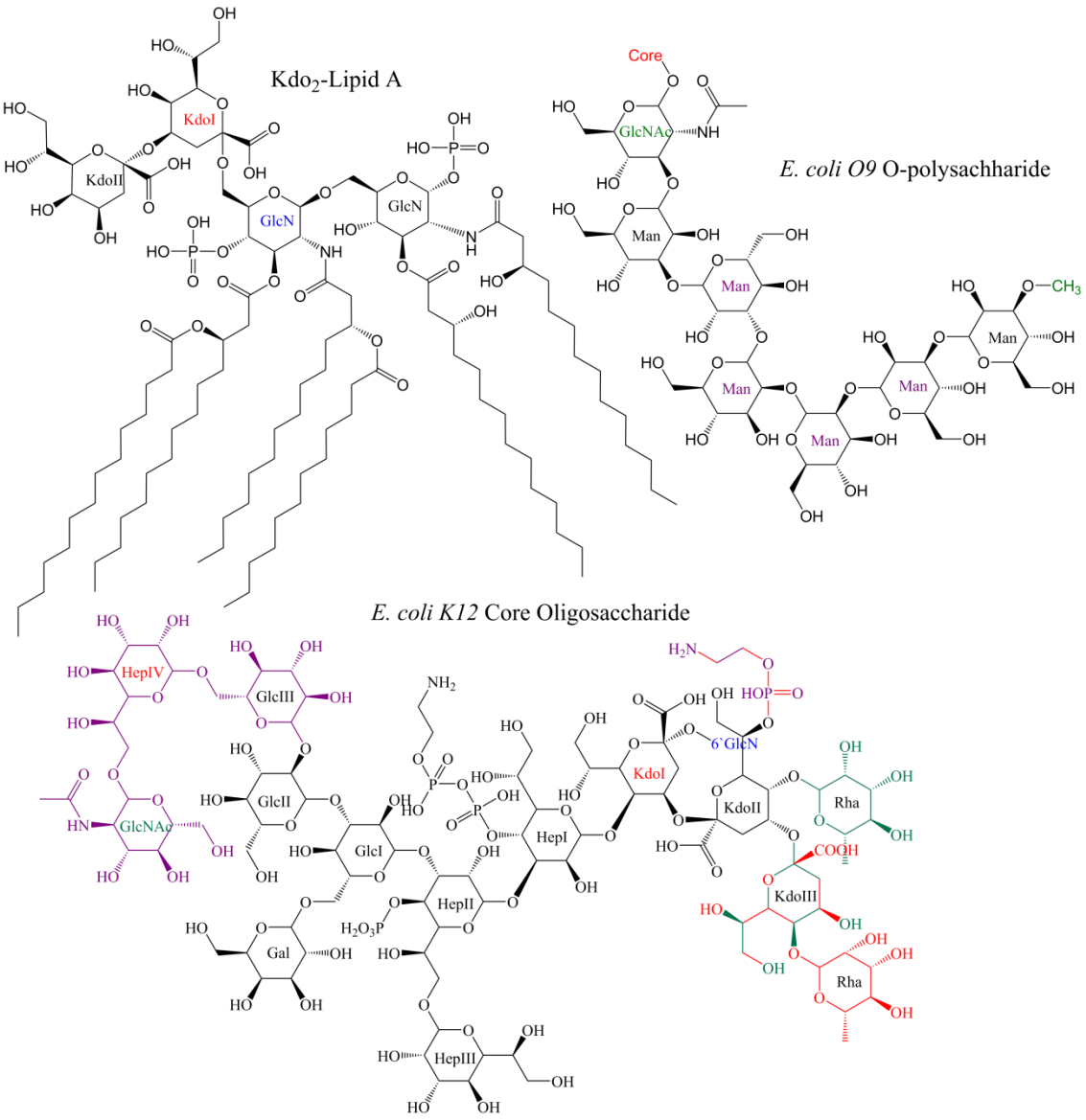


Figure 1.1: LPS chemical structure. Chemical structures of the canonical lipid A of *E. coli*, the core oligosaccharide of *E. coli* K-12, and the *E. coli* O9 O-antigen (showing one repeat unit) (1). The text for the distal glucosamine of lipid A attached to the core oligosaccharide is colored blue. The text for the residues at each end of the core oligosaccharide is colored red. The text for the ends of the O-antigen is colored green, and that for the repeat region is colored purple. In the core oligosaccharide, atom and bond coloring indicates different core structures (11). Black indicates the conserved structure, and purple, red, and green indicate moieties present in three different variants of the K-12 core. Note that other K-12 core structures have also been characterized (12). Dual coloring of moieties indicates their presence in both variants.

CD14, which may be secreted or lipid-linked to the extracellular face of cells, transfers this LPS to the complex of myeloid differentiation factor 2 and Toll-like receptor 4 (MD2/TLR4) (9, 13–15). Activation of TLR4 by LPS leads to the production and secretion of pro-inflammatory cytokines and type I interferon (9, 16). While this response is beneficial for clearing small bacterial infections, overstimulation of the inflammatory response during systemic infection, such as sepsis (presence of bacteria in the blood), is potentially fatal and can cause permanent organ damage or neurological problems (17, 18). Because of the severity of this condition, rapid treatment with broad-spectrum antibiotics before the causative bacteria can be identified is recommended for patients with sepsis (17). The highly conserved and (in most cases) essential enzymes involved in the synthesis of the lipid A portion of LPS are promising targets for the development of new broad-spectrum antibiotics to treat sepsis and infections by antimicrobial resistant Gram-negative strains (1, 5, 7, 19).

Extensive biochemical studies have described the contribution of the different moieties of lipid A to the induction of the inflammatory response by LPS (20–23). The number and position of acyl chains is the most important factor in determining inflammatory potential in humans (20–23). Canonical hexa-acylated lipid A from *E. coli* (Figure 1.1) induced interleukin-1 production at concentrations as low as 10 pg/mL (20). Removal of the secondary acyl chain (an acyl chain attached to an acyl chain directly linked to glucosamine) from the distal 3'-position increased the concentration required for induction 100-fold, and addition of an acyl chain to the proximal 2-position also increased active concentration 100-fold (20, 22). These effects were multiplicative: a hexa-acylated lipid A derivative with secondary chains at the 2/2'-positions of both

proximal and distal glucosamines only induced the inflammatory response at 10,000-fold higher concentration (20, 22). Little activity was observed with a penta-acylated derivative with the only secondary chain at the proximal 2-position, and lipid A derivatives with 4 or fewer chains showed no inflammatory induction (20–22). In fact, tetra-acylated lipid A derivatives could competitively inhibit the inflammatory activity of lipid A or LPS (20–23). Interestingly, this effect is not conserved among mammals as tetra-acylated lipid A derivatives retained inflammatory activity in mice (21, 23). In addition, the 1- and 4'-phosphates are important for induction of inflammation (20–22). Removal of one of the phosphates increased the active concentration 100-fold (20, 22). Replacing the 1-phosphate with a phosphonooxyethyl group had little effect, but changing the position of the phosphonooxyethyl group from the axial to the equatorial position (changing from the α- to the β-anomer of the proximal glucosamine) increased the minimum active concentration to 1 μg/mL (21, 22).

The structural basis of the inflammatory potential of differentially substituted lipid A moieties has been revealed by crystal structures of LPS and tetra-acylated lipid A derivatives (competitive inhibitors) bound to MD2 and the MD2/TLR4 complex (13, 14, 24). The LPS-bound complex was crystallized as a dimer of MD2/TLR4-extracellular domain heterodimers (Figure 1.2) (14). MD2 forms a β-cup fold wherein two β-sheets are separated to form a large, hydrophobic pocket that accommodates 5 of the 6 acyl chains of lipid A (14, 24). The acyl chain in the proximal 2-position binds outside of the MD2 pocket and interacts with the TLR4 subunit of the other heterodimer (Figure 1.2) (14).

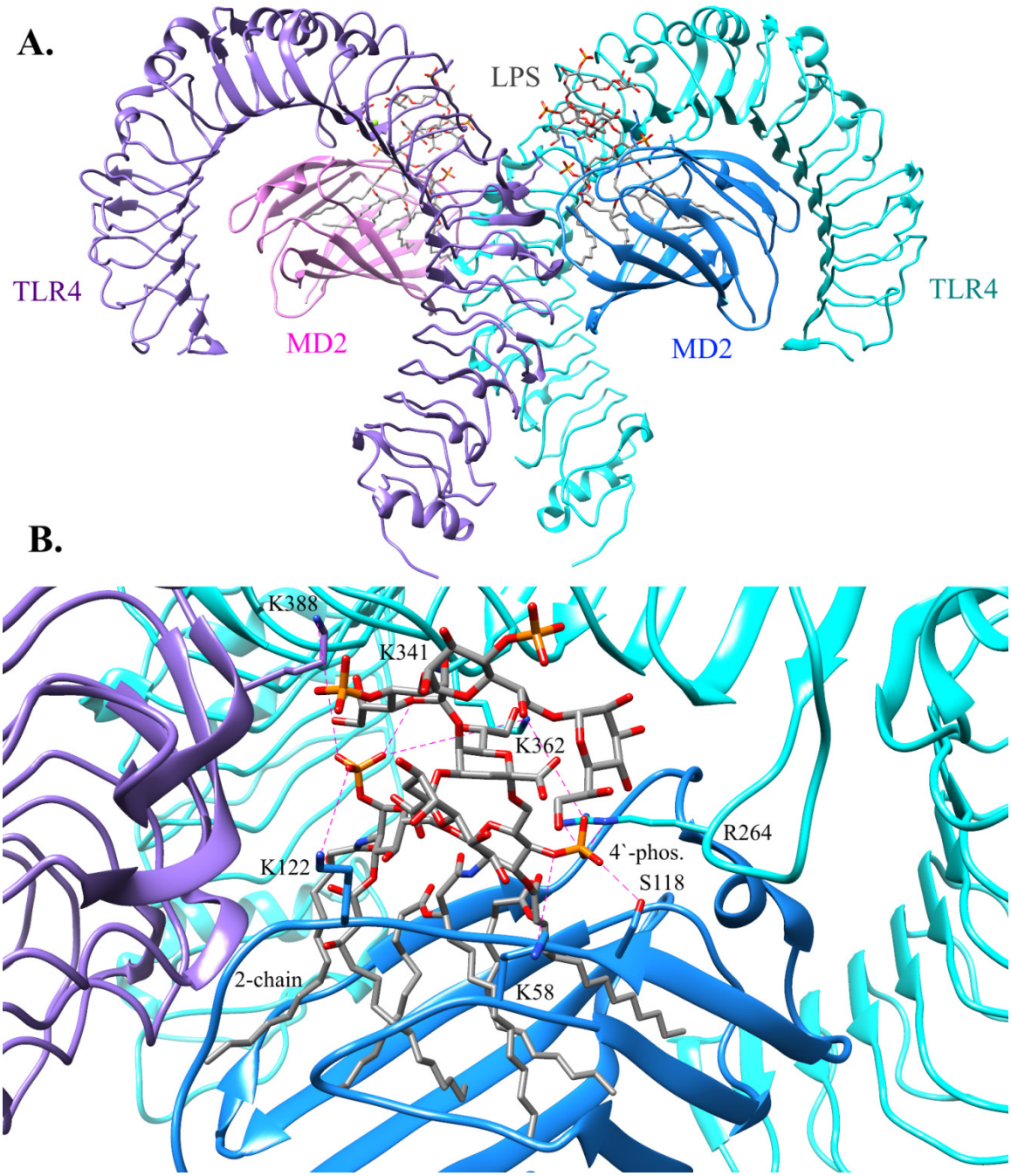


Figure 1.2: TLR4/MD2 complex with LPS. A. Heterotetramer of TLR4 and MD2 bound to LPS with the core oligosaccharide resolved out to the third Hep (PDB: 3FXI) (14). **B.** Zoom-in of one LPS in **A.** showing long distance electrostatic and hydrogen-bonding interactions. Distances of interactions in Ångstroms are as follows: 9.0 from 1-phosphate O2 to $K122_C$ Nζ, 7.3 from 1-phosphate O4 to $K388_B$ Nζ, 5.6 from 1-phosphate O3 to $K341_A$ Nζ, 8.3 from 1-phosphate O3 to $K362_A$ Nζ, 3.0 from 4'-phosphate O2 to $S118_C$ Oγ, 3.4 from 4'-phosphate O3 to $R264_A$ Nη2, 3.3 from 4'-phosphate O3 to $R264_A$ Nε, 10.4 from distal

glucosamine 4'O to $K58_C$ ζ, and 7.6 from 4'-phosphate O4 to $K362_A$ Nζ. Figure generation and structure analysis were performed in UCSF Chimera (25).

The base glucosamine disaccharide is bound by MD2 while the inner core oligosaccharide, consisting of two 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) residues and three L-glycero-D-mannoheptose (Hep) residues, interacts with the TLR4 of the same heterodimer, and the remaining sugar residues extend into solvent and are not visible (14). The 1-phosphate forms long distance (>4 Å) electrostatic interactions with K122 of MD2, K341 and K362 of the TLR4 in the same heterodimer, and K388 of the other TLR4 subunit (Figure 1.2) (14). The 4'-phosphate forms a hydrogen-bond with S118 of MD2, forms a salt-bridge with R264 of the TLR4 in the same heterodimer, and forms long distance electrostatic interactions with K58 of MD2 and K362 of the TLR4 in the same heterodimer (14). However, when the MD2/TLR4 complex was bound by the tetra-acylated lipid A derivative and competitive inhibitor Eritoran, all 4 acyl chains, including the chain at the proximal 2-position, were bound in the pocket MD2 (24). The tetra-acylated precursor of lipid A (lipid IV_A) bound similarly to MD2 (13). Furthermore, Eritoran bound more deeply in the pocket precluding interactions between the phosphates and TLR4 (14, 24). This structure was crystallized as a heterodimer, and the lack of heterotetramer formation was supported by solution behavior of the Eritoran-bound MD2/TLR4 complex on a gel filtration column (24). On the other hand, the addition of LPS led to the formation of an oligomer consistent with a heterotetramer (24). As oligomerization of TLR4 is thought to be required for activation of downstream signaling, these structures provide a simple explanation for how tetra-acylated lipid A derivatives inhibit induction of inflammation by lipid A/LPS: tetra-acylated lipid A

derivatives occupy the MD2 pocket without triggering heterotetramer formation and downstream signaling and thus block binding of lipid A molecules that can trigger downstream signaling (9, 13, 14, 24).

However, interactions between LPS and the host's immune system are not limited to the inflammatory response. In addition, modifications to the acylation and charge of LPS can protect bacteria from being killed by cationic antimicrobial peptides (CAMPs) and the complement system, which both act to disrupt bacterial membranes (3, 8, 18, 26–29). CAMPs act by interacting with anionic groups in LPS, particularly the lipid A phosphates, and inserting a hydrophobic tail into the membrane to disrupt the permeability barrier (8, 18). The complement system acts by antibody-targeted, lectin-targeted, or spontaneous binding of complement system factors to the bacterial surface that ultimately leads to the insertion of pore-forming proteins into the bacterial membrane (26). The bacterial outer membrane protein (PagP) adds a secondary palmitate to the β-hydroxyl group of the proximal 2-acyl chain and sometimes, as in *Bordetella parapertussis*, also adds palmitate to an available β-hydroxyl group of the distal 3'-acyl chain (8, 30, 31). PagP activity has been found to increase resistance to some CAMPs and the complement system (30, 31). In addition, PagP activity has been linked to biofilm formation in several bacterial species and to *E. coli* biofilm persistence in rats (32, 33). Another bacterial outer membrane protein (PagL) removes the acyl chain from the 3-hydroxyl of the proximal glucosamine (8, 29). PagL was found to contribute to the resistance to polymyxin B (a CAMP) in a resistant strain of *Pseudomonas aeruginosa*, and neutron reflectometry indicated that polymyxin B failed to penetrate into a bilayer with PagL-modified lipid A (29). Addition of cationic moieties to mask the negative

charge of LPS can also decrease susceptibility to CAMPs and the complement system (6, 8, 18, 28, 33–35). EptA transfer of phosphoethanolamine to the lipid A phosphates increased resistance to polymyxin B (8, 35). In addition, the periplasmic transferase ArnT transfers 4-amino-4-deoxy-L-arabinose (L-Ara4N) to the lipid A 4'-phosphate, and addition of this positively charged sugar increased resistance to the CAMPs polymyxin B and LL-37 (8, 33, 36). Finally, changes in the glycosylation of the core oligosaccharide or O-antigen can help bacteria evade the host's immune system (3, 18, 26). In *Helicobacter pylori*, the O-antigen can be modified with fucose to resemble host antigens, and in *Pseudomonas aeruginosa*, the O-antigen may be completely lost during chronic infections (3). Likewise, *Neisseria meningitidis* can alter its core oligosaccharide to resemble an erythrocyte antigen, or *N. meningitidis* can add sialic acid to its core oligosaccharide to recruit host factor H, which inhibits killing by the complement system (18, 26).

1.2 Determination of the Structure of Lipid A

Research describing the relevance of lipid A moieties to induction of inflammation could not have been accomplished prior to the elucidation of the structure of lipid A. This section describes how the structure of Kdo_2-lipid A presented in Figure 1.1 was determined.

Some of the earliest research into the structure of lipid A identified the glucosamine disaccharide as the central component, and the glycosidic linkage was assigned as β1-6 on the basis of reactivity of the N-acetylated disaccharide with p-

dimethylaminobenzaldehyde (in Ehrlich reagent) and non-stereospecific and β-linkage-specific acetylglucosaminidases (37). The identity of the central disaccharide was further supported by comparison with β1-6 and β1-4 glucosamine disaccharide standards by thin layer chromatography (TLC) (38). Nuclear magnetic resonance (NMR) of Kdo-lipid A showed that ^{13}C chemical shifts and J_{CH} coupling constants were consistent with the α-anomer of the proximal (reducing end) glucosamine and the β-anomer of the distal glucosamine (39). Finally, NMR of lipid A precursors showed a change in the chemical shift of C6 of the proximal glucosamine consistent with formation of the 1-6 linkage and showed that ^{1}H chemical shifts and $J_{1,2}$ coupling constants were also consistent with the above anomer assignments (40, 41).

In addition, early research established the positions of the phosphate groups (37, 42). Protection of the reducing end of the disaccharide by an acid labile phosphate established the presence of phosphate attached to the anomeric carbon of the proximal glucosamine (37). Protection of the distal glucosamine by a less labile phosphate group from periodate cleavage in the presence (but not absence) of N-acetylation indicated the presence of phosphate at the 4'-position (37, 42). ^{31}P NMR showed that 1-diphosphate is sometimes present on the proximal glucosamine (39).

Determination of the positions of the acyl chain substituents required the combination of several studies performed over several years. Because amides are less labile than esters, the β-hydroxymyristate at the 2- and 2'-amino groups of the glucosamine residues were the first acyl chains located (43). Selective removal of these acyl chains by fatty acyl amidases further demonstrated that β-hydroxymyristate is present at the amino group of both glucosamines (38). Iodomethane-based cleavage of amides followed by analysis

of released fatty acids by gas-liquid chromatography showed the presence of β-hydroxymyristate substituted at the β-hydroxyl group with a secondary fatty acid chain, primarily laurate, myristate, and palmitate (44). More selective release of laurate and palmitate substituted β-hydroxymyristate by iodomethane suggested that these were released from the 2- and 2'-amino groups while myristate substituted β-hydroxymyristate was ester-linked to glucosamine (44). A breakthrough in the understanding of the structure of lipid A was brought about by the fortuitous discovery of the lipid A precursor lipid X (45). Characterization of lipid X by mild alkaline hydrolysis and ^{1}H NMR located β-hydroxymyristate esterified at position 3 in addition to the amide-linked β-hydroxymyristate at position 2 (46, 47). This led Takayama *et al*. (47) to propose a semi-symmetrical model of lipid A with β-hydroxymyristate or β-acyloxymyristate at the 2-, 3-, 2'-, and 3'-positions of the glucosamine disaccharide and with the Kdo residues attached at the 6'-hydroxyl. This was an important change to the lipid A model because it correctly described the sites of acylation and the attachment site of the core oligosaccharide, which was previously placed at the 3'-position (42, 47). Selective ester hydrolysis in combination with mass spectrometry showed that the laurate and myristate substituted β-hydroxymyristates both reside on the distal glucosamine of lipid A and that the proximal glucosamine has one amide-linked and one ester-linked β-hydroxymyristate (48, 49). Selective ester hydrolysis in combination with ^{13}C NMR confirmed that the acyl groups are attached at the 2-, 3-, 2'-, and 3'-positions of the glucosamine disaccharide in Kdo-lipid A and showed that one of the two amide-linked β-hydroxymyristates is substituted with a normal fatty acid (50). Thus, the positions of the acyl chains shown in Figure 1.1 were established. However, as mentioned above, palmitate substituted β-

hydroxymyristate was also a major component of lipid A (44). Mass spectrometry in addition to selective ester hydrolysis placed this moiety at the 2-position of the proximal glucosamine, but mass spectrometry also showed that this modification is present in *Salmonella* but not *Escherichia* under optimal growth conditions (49, 51).

Finally, early research into the structure of lipid A determined that 2 or 3 Kdo residues are attached to the glucosamine disaccharide (37). As shown in Figure 1.1, it is possible for 3 Kdo residues to be present in the inner core oligosaccharide; however, the inner core is more variable between bacterial species than lipid A (1). Two relatively well conserved Kdo residues are added early in LPS synthesis (1, 52, 53). As stated above, Kdo was initially thought to be attached at the 3'-hydroxyl of the glucosamine disaccharide, but this was later suggested to be a site of acylation (42, 46). ^{13}C NMR chemical shifts and integrations and ^{1}H NMR chemical shifts from multiple studies were consistent with the attachment of 2 Kdo residues at the 6'-hydroxyl of lipid A (39, 50, 54, 55). ^{13}C NMR and ^{1}H NMR chemical shifts and J_{CH} coupling constants indicated that both Kdo residues are α-anomers with an α2-4 glycosidic bond linking them and an α2-6 bond between the first Kdo and the distal glucosamine (50, 55). The α2-4 glycosidic bond was further corroborated by gas-liquid chromatography-tandem mass spectrometry (56).

1.3 Raetz Pathway

The canonical Raetz pathway of lipid A synthesis is a 9 enzyme pathway that produces Kdo_2-lipid A (Figure 1.3) (1). While some variation exists between species, particularly with regards to acylation, the Raetz pathway, especially the first 4 enzymes,

is well conserved in Gram-negative bacteria, excluding specialized species that do not produce LPS (7, 8). Most of these enzymes have been characterized at the biochemical and structural level, and the following summarizes what is known about the enzymes of the Raetz pathway.

LpxA

Early experiments with radiolabeled substrates indicated that UDP-N-acetylglucosamine is incorporated into the lipid A synthesis pathway and suggested that acylation of the 3-hydroxyl of glucosamine precedes the removal of acetate (57). The acyltransferase responsible for the addition of this acyl chain was denoted LpxA and was found to be specific for UDP-N-acetylglucosamine and R-3-hydroxymyristoyl-acyl carrier protein (ACP), with little activity toward the S-enantiomer or the coenzyme A adduct of β-hydroxymyristate or toward palmitoyl-ACP, myristoyl-ACP, or acyl chain acceptors with chains more than 3 carbons long at the 2-amine of glucosamine or with uracil substituted with any base other than thymine (58, 59). Furthermore, radiolabeled LpxA product was converted to other intermediates of lipid A synthesis (58).

Crystal structures of LpxA revealed that LpxA is composed of an unusual N-terminal left-handed parallel β-helix domain and a C-terminal α-helical domain and forms a trimer with 3-fold symmetry (Figure 1.4) (60). Each β-helix approximates an equilateral triangular prism, and these helices come together in the trimer to form a large cleft between each pair of subunits (60). Extended loops within helical repeats 4 and 5 form additional contacts between adjacent subunits (61).

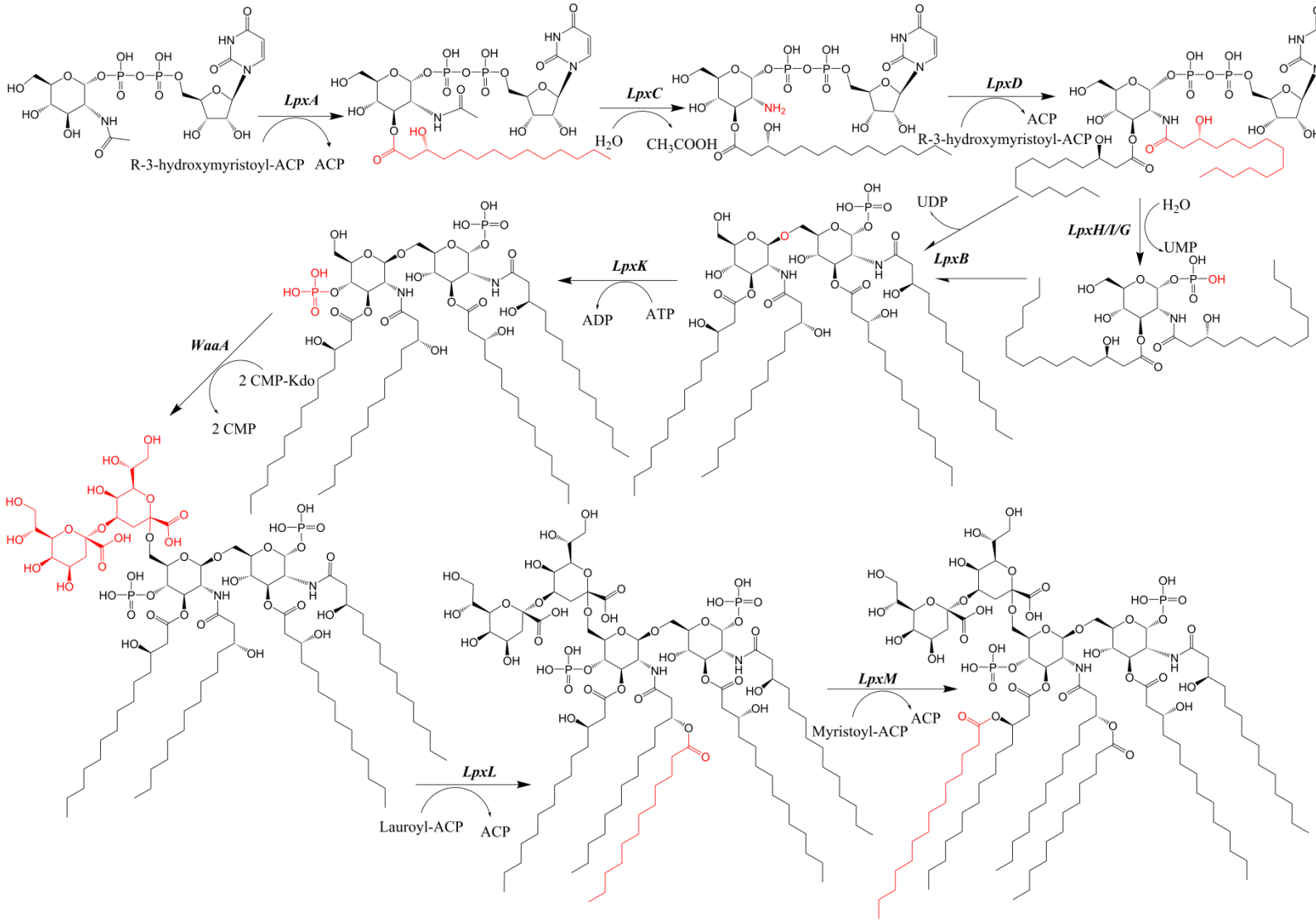

LpxA
R-3-hydroxymyristoyl-ACP
ACP
LpxC
H2O
CH3COOH
LpxD
R-3-hydroxymyristoyl-ACP
ACP
UDP
LpxB
H2O
LpxH/I/G
UMP
LpxK
ADP
ATP
WaaA
2 CMP-Kdo
2 CMP
LpxL
Lauroyl-ACP
ACP
LpxM
Myristoyl-ACP
ACP

Figure 1.3: Raetz pathway of lipid A synthesis. The moieties added in each step are shown in red (1).

Crystal structures with UDP-N-acetylglucosamine or the product bound show that substrates bind in this cleft between subunits with the uridine and N-acetylglucosamine moieties contacting adjacent subunits and with the R-3-hydroxymyristoyl chain extending up the cleft toward the C-terminal domain (Figure 1.4) (61, 62). The extended loop within helical repeat 4 provides part of the contacts for N-acetylglucosamine (61, 62).

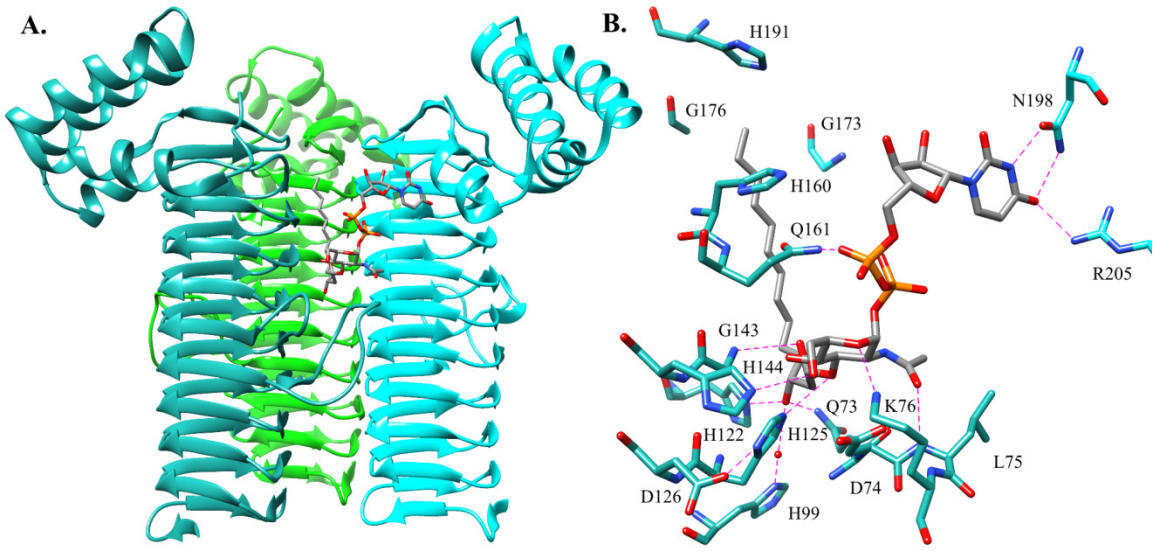


Figure 1.4: LpxA. A. Ribbon diagram of the LpxA trimer bound to its 3-O-(R3-hydroxymyristoyl)-N-acetylglucosamine product (PDB: 2QIA) (62). **B.** Residues involved in product binding from **A.** Subscripts are used below to distinguish subunits in the trimer. Hydrogen-bonding interactions are shown. Distances of interactions in Ångstroms are as follows: 3.3 from 3'-acyl carbonyl O to $G143_{0.3}$ N, 3.2 from uracil O4 to $N198_{0.1}$ Nδ, 3.0 from uracil O4 to $R205_{0.1}$ Nη2 2.7 from uracil N3 to $N198_{0.1}$ Oδ, 2.9 from water to $H99_{0.3}$ Nδ, 2.8 from same water to β-hydroxyl, 2.8 from β-hydroxyl to $H122_{0.3}$ Nε, 3.0 from β-hydroxyl to $Q73_{0.3}$ Nε, 3.3 from glucosamine 3'O to $H125_{0.3}$, 2.8 from $H125_{0.3}$ Nδ to $D126_{0.3}$, 2.9 from glucosamine 6'-hydroxyl to $H144_{0.3}$ Nε, 3.0 from glucosamine ring O to $K76_{0.3}$ Nζ, 3.0 from 2'-acetyl carbonyl to $L75_{0.3}$ N, and 2.6 from α-phosphate to $Q161_{0.3}$ Nε.

Furthermore, these crystal structures provide insight into the mechanism and selectivity of LpxA (61, 62). LpxA is thought to catalyze the nucleophilic attack of the

glucosamine 3-hydroxyl on the thioester of R-3-hydroxymyristoyl-ACP, and a conserved His (H125 in *E. coli* LpxA) is positioned to act as the catalytic base and accept the proton from this hydroxyl while the backbone amide of G143 is positioned to stabilize the oxyanion intermediate (61, 62). The role of H125 was supported by activity assays that showed it to be critical for activity (63). D126 hydrogen-bonds with and stabilizes the position of H125 forming a catalytic dyad similar to that observed in other acyltransferases (61). In addition, the product-bound structure suggests that H191 of *E. coli* LpxA limits the length of the acyl chain by capping the channel wherein the chain binds, and mutation of another residue that lines this channel (G173) to larger residues decreased activity toward the natural acyl donor substrate and increased activity toward the 10 carbon chain analogue (62, 64). Moreover, selection of the R-3-hydroxylated fatty acid may be controlled by hydrogen-bonding interactions between this functional group and H122, Q73, and an ordered water bound to H99 (62). Finally, UDP-N-acetylglucosamine selectivity is governed by extensive hydrogen-bonding as well as the stereospecific architecture of the N-acetylglucosamine binding pocket that would interfere with the binding of a sugar with axial substituents (Figure 1.4) (61).

LpxC

LpxC catalyzes the second step in the Raetz pathway, the deacetylation of UDP-3-O-(R3-hydroxymyristoyl)-N-acetylglucosamine (65, 66). Early experiments showed that the LpxA product is deacetylated when incubated with *E. coli* extracts and that the deacetylated product can be converted into other intermediates in the pathway (65). Furthermore, LpxC was found to have poor activity toward UDP-N-acetylglucosamine (67). Because the reaction catalyzed by LpxA is freely reversible and, surprisingly,

thermodynamically unfavorable, LpxC catalyzes the committed step of lipid A synthesis (59). In accordance with the thermodynamic importance of this step, LpxC is also an important point of regulation for lipid A synthesis (68, 69). LpxC activity is regulated by proteolytic degradation by FtsH, which may be controlled by the levels of acyl-ACP or Raetz pathway intermediates (68, 69).

LpxC is a Zn^{2+}-dependent metalloenzyme though it displays higher activity in the presence of Ni^{2+} or Co^{2+}, which may be due to inhibition by Zn^{2+} binding at a second site (67). Consistent with these enzymatic assays, the crystal structure of *Aquifex aeolicus* LpxC was solved with two Zn^{2+} in the active site, representing the zinc-inhibited state, and the crystal structure of *E. coli* LpxC in complex with the lipidic product, which was crystallized in a high concentration of phosphate, was solved with a single high-affinity Zn^{2+} in the active site (70, 71). Native mass spectrometry was also consistent with the binding of one high-affinity Zn^{2+} (71).

Solution NMR and crystal structures of LpxC revealed that the enzyme is composed of two α/β-domains with a conserved topology that form a five-stranded β-sheet and two α-helices (70, 72). The α-helices of the conserved fold are packed together giving LpxC an overall pseudo-two-fold symmetry (Figure 1.5) (70, 72). The domains contain divergent insertions between the fourth β-strand and first α-helix: the N-terminal domain's insertion forms a three-stranded β-sheet, and the C-terminal domain's insertion contributes a short helix to the helical core and forms an α/β-subdomain (70–72).

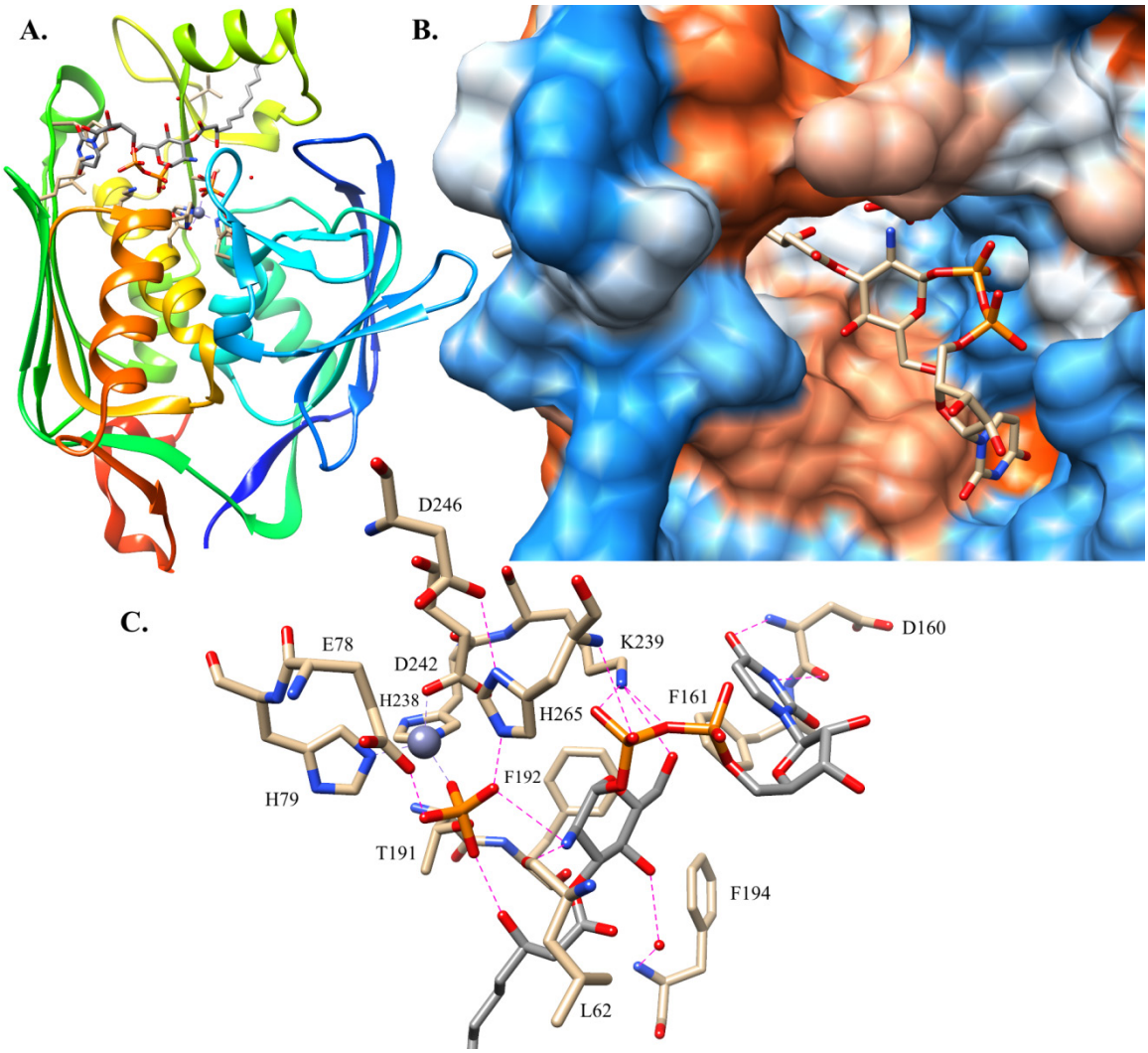


Figure 1.5: LpxC. A. Ribbon structure of LpxC bound to its UDP-3-O-(R3-hydroxymyristoyl)-glucosamine product (PDB: 4MDT) (71). Spectrum coloring begins with blue at the N-terminus. **B.** Hydrophobicity surface of structure in **A.** with orange-blue scale. Orange is more hydrophobic. **C.** Residues involved in product binding from **A.** Zinc coordination is shown in purple, and hydrogen-bond and salt-bridge interactions are shown in pink. Distances of interactions in Ångstroms are as follows: 1.9 from Zn^{2+} to H238 Nε, 2.0 from Zn^{2+} to H79 Nε, 2.0 from Zn^{2+} to D242 Oδ1, 1.9 from Zn^{2+} to phosphate O2, 2.5 from water to β-phosphate and 3.2 to α-phosphate and 3.0 to A266 N, 3.2 from pyrophosphate bridging O and 2.8 from β-phosphate and 3.5 from glucosamine 6'-hydroxyl to K239 Nζ, 2.8 from β-hydroxyl to phosphate O1, 3.2 from water to F194 N and 2.9 to glucosamine 4'-hydroxyl, 3.1 from glucosamine 2'-amine to phosphate O3 and 2.7 to L62 O, 2.6 from E78 Oε2 to phosphate O4,2.4 from H265 Nε to phosphate O3, 3.5 from uracil O4 to D160 N, 2.7 from D246 Oδ2 to H265 Nδ, and 3.5 from uracil N3 to D160 O.

The structure of *E. coli* LpxC with UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine and phosphate bound in the active site shows that the product primarily binds to the insertions and the C-terminal domain with the first helix of the N-terminal domain contributing one Zn^{2+}-coordinating residue and the putative catalytic base (71). Zn^{2+} is coordinated with tetrahedral geometry by two His side chains, one Asp side chain (H79, H238, and D242 in *E. coli* LpxC), and a molecule from the crystallization solution (70–72). The importance of these residues was supported by enzymatic assays that showed decreased activity in mutants of any one of these residues as well as greatly decreased Zn^{2+}-binding in the His to Ala mutants (73).

In the product-bound structure of *E. coli* LpxC, the fourth ligand of Zn^{2+} was modeled as phosphate, and it may approximate the tetrahedral transition state of the leaving group, hydrogen-bonding with the glucosamine amine (71). The phosphate forms hydrogen-bonds with E78, T191, and H265 implicating these residues in catalysis, and mutants of these residues were indeed found to have decreased activity (71, 73, 74). Specifically, E78 is positioned to accept a proton from a Zn^{2+}-coordinated hydrolytic water, and this role was supported by the greater reduction in activity of E78Q in comparison to E78A as E78Q cannot act as a catalytic base but likely blocks the compensatory positioning of another residue or solvent molecule (71, 73). On the other hand, H265 is positioned to donate a proton to the glucosamine amine, and D246, which hydrogen-bonds with the opposite nitrogen of the H265 azole thus tuning the pK_a of H265, was also critical for activity (71, 73).

As stated above, the product primarily interacts with the insertions and the C-terminal domain (71). The UDP moiety binds to the C-terminal domain: uracil hydrogen-

bonds with the backbone of D160 and forms a π-stacking interaction with F161, and the pyrophosphate interacts with K239 and the backbone of H265 (Figure 1.5) (71). Recognition of the glucosamine moiety involves all three subdomains (71). The 6'-hydroxyl also hydrogen-bonds with K239, and the 2'-amine hydrogen-bonds with the backbone of L62 (Figure 1.5). The importance of K239 was confirmed by enzymatic assays showing a slightly increased K_M and greatly decreased k_{cat} in the K239A mutant (74). In addition, the 4'-hydroxyl hydrogen-bonds to an ordered water bound to the F194 backbone (71). Finally, the acyl chain binds in a hydrophobic channel formed by the C-terminal domain insertion with the chain terminus extending into solvent, suggesting that, while presence of the chain is critical for activity, LpxC has little specificity for chain length (Figure 1.5) (67, 70, 71).

LpxD

Early experiments indicated that addition of the second R3-hydroxymyristoyl group at the glucosamine 2-amino group precedes the hydrolysis of the phospho-anhydride of UDP (57). This structural gene for the UDP-3-O-(R3-hydroxymyristoyl)-glucosamine N-acyltransferase (*lpxD*) was identified by homology to *lpxA* and its position in the same operon, and enzymatic assays showed LpxD is highly specific for β-hydroxymyristoyl-ACP, showing no activity for myristoyl-ACP and 10-fold reduced activity for β-hydroxylauroyl-ACP (75, 76). Furthermore, steady-state kinetic experiments showed that LpxD follows an ordered sequential mechanism wherein R3-hydroxymyristoyl-ACP binds first, and ACP dissociates last (76).

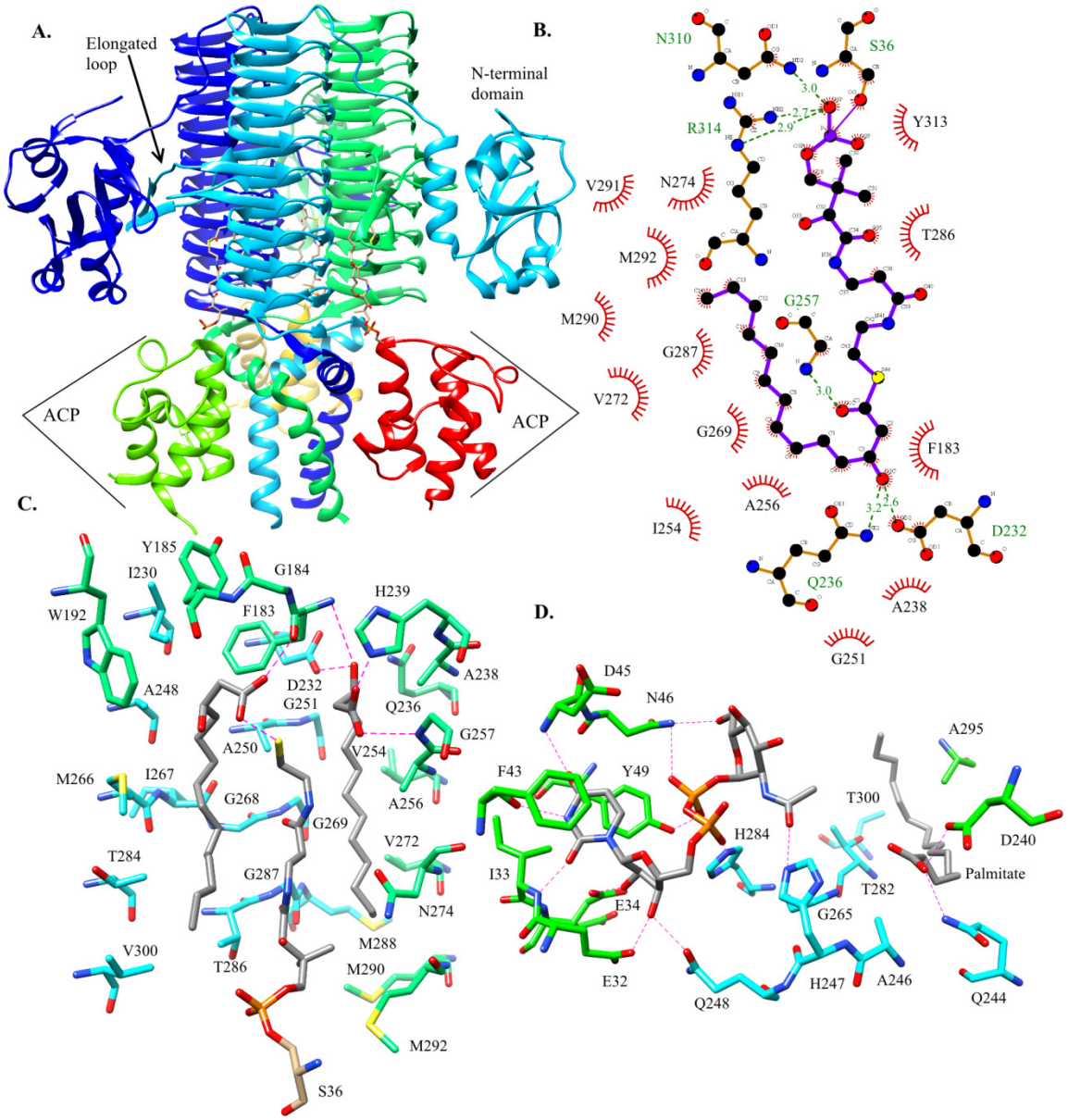


Figure 1.6: LpxD. A. Heterohexamer of LpxD and R3-hydroxymyristoyl-ACP (PDB: 4IHF) (77). ACP chains are colored red, yellow, and chartreuse, and LpxD chains are colored blue, cyan, and sea green. **B.** LigPlot+ diagram of acyl donor substrate from **A.** (78). Distances of hydrogen-bond and salt-bridge interactions shown in Ångstroms are as follows: 3.2 from acyl chain β-hydroxyl to Q236 Nε and 2.6 to D232 Oδ2, 3.0 from acyl chain carbonyl to G257 N, 3.0 from phosphate to N310 Nδ and 2.7 and 2.9 respectively to R314 Nη and Nε. **C.** Acyl chain binding in LpxD/ACP heterohexamer (PDB: 4IHG) (77). Subscripts are used below to distinguish protein chains in the complex. Hydrogen-bonds are shown, corresponding distances in Ångstroms are as follows: 2.9 from putative 2'-R3-hydroxymyristate carboxylate to $G257_C$ N, 2.1 from putative 2'-R3-hydroxymyristate carboxylate to $H239_C$ Nε, 3.0 from putative 2'-R3-hydroxymyristate β-hydroxyl to $D232_B$ Oδ2 and 3.6 to $F183_C$ N, and 2.8 from putative 3'-

R3-hydroxymyristate carboxylate to $F183_C$ O and 3.2 to S of phosphopantetheine attached to ACP $S36_H$. **D.** UDP-N-acetylglucosamine and palmitate binding to LpxD trimer (PDB: 2IU9) (79). Hydrogen-bonds are shown with the following distances in Ångstroms: 2.8 from palmitate carboxylate to $Q244_A$ Nε and 2.9 to $D240_B$ Oδ2, 2.6 from 2'-acetyl carbonyl to $H247_A$ Nε, 2.8 from uracil O2 to $I33_B$ N, 2.7 from uracil N3 to $F43_B$ O, 2.8 from uracil O4 to $D45_B$ N, 2.6 from ribose 3'-hydroxyl to $E32_B$ Oε2 and 2.9 to $Q248_A$ Oε, 3.0 from ribose 2'-hydroxyl to $E34_B$ Oε1, 2.4 from β-phosphate to $Y49_B$ Oη and 3.2 to $N46_B$ Nδ, 2.7 from α-phosphate to $H284_A$ Nε, and 3.1 from glucosamine 6'-hydroxyl to $N46_B$ Nδ.

Crystal structures of LpxD reveal many similar features shared with LpxA (79, 80). In particular, LpxD has a very similar left-handed β-helix domain that includes two elongated loops (within helical repeats 5 and 6) that contribute additional interactions with the adjacent subunit in the trimer (79). Moreover, LpxD also has an α-helical C-terminal domain; though in LpxD, the C-terminal helices of the subunits come together to form a trimeric helix bundle (79). However, LpxD has an additional N-terminal domain composed of a five-stranded β-sheet surrounded by α-helices and a two-stranded β-sheet (Figure 1.6) (79). Furthermore, crystal structures of LpxD bound to substrates and substrate analogues have revealed conserved catalytic residues and mechanisms of substrate specificity (77, 79). Specifically, a conserved His (H239 in *E. coli*) is positioned to accept a proton from the glucosamine amine while the G257 amide is positioned to stabilize the oxyanion intermediate (77, 79, 80). As in LpxA, the substrates bind at the cleft between subunits (77, 79).

In the structure of *Chlamydia trachomatis* LpxD, UDP-N-acetylglucosamine binds to the N-terminal domain and the left-handed β-helix domain of the adjacent subunit with the N-terminal domain primarily binding the uridine moiety (79). At the A+B chain active site, uracil is sandwiched in π-stacking interactions with $F43_B$ and $Y49_B$ (F41 and Y47 in *E. coli*) and forms hydrogen-bonds with the backbone of residues

33_B and 43_B, and 45_B (Figure 1.6) (79). Ribose is hydrogen-bonded to $E32_B$ and $E34_B$ (S30 and Q32 in *E. coli*) and $Q248_A$ (Q240 in *E. coli*) while the pyrophosphate is engaged by $Y49_B$, $N46_B$, and $H284_A$ (N44 and H276 in *E. coli*), and N-acetylglucosamine binds to $N46_B$ and $H247_A$ (H239 in *E. coli*, the predicted catalytic base) (79). Kinetic assays in *E. coli* LpxD confirmed that H239 was critical for catalysis: H239A decreased k_{cat} almost 3 orders of magnitude with a moderate decrease of acyl acceptor K_M (76). In addition, F41 was confirmed to be critical for acyl acceptor binding: F41A increased acceptor K_M 29-fold with only a small decrease in k_{cat} (76). Conversely, Y47A had little effect, and H276A decreased k_{cat} 32-fold (76).

E. coli LpxD crystallized in complex with ACP yielded structures with R3-hydroxymyristoyl-ACP and with holo-ACP and two β-hydroxymyristate molecules bound to LpxD (77). ACP primarily binds to the C-terminal domain with the pantathionine arm on ACP S36 extending up the active site cleft (Figure 1.6) (77). Association with ACP is largely electrostatic with an acidic patch on ACP contacting a basic patch on LpxD (77). For example, the Ala mutant of R193, which forms an ion pair with E41 of ACP, was found to cause a 23-fold increase of the R3-hydroxymyristoyl-ACP K_M (76, 77). In the intact substrate structure, the R3-hydroxymyristoyl chain turns 180^o and extends back toward the C-terminus of LpxD deeper in the cleft (77). Specificity for the length of the acyl chain attached to ACP is conferred by the length of the hydrophobic pocket that accommodates the chain (76, 77, 79). In *E. coli* LpxD, the ~13 Å deep pocket is terminated by M290, and M290A allowed incorporation of 16 and 18 carbon acyl chains *in vivo* (80). Moreover, in *C. trachomatis* LpxD, which transfers a β-hydroxyarachidoyl chain, the corresponding residue is G298, and the hydrophobic

pocket is ~18 Å deep (79). In the B+C active site, the β-hydroxyl group of the transferred chain is recognized by hydrogen-bonds to $D232_B$ and $Q236_C$ of *E. coli* LpxD, and the acyl chain carboxylate hydrogen-bonds to the $G257_C$ amide, consistent with its proposed stabilization of the oxyanion (Figure 1.6) (76, 77, 80). Finally, the structure of *E. coli* LpxD bound to holo-ACP and two β-hydroxymyristates reveals the binding pocket for the 3'-acyl chain of the acceptor substrate: this pocket is formed by the left-handed β-helix of the subunit that contributes the N-terminal domain, the elongated loop of repeat 5 from the adjacent subunit, and the pantathionine group (Figure 1.6), consistent with kinetic experiments that showed R3-hydroxymyristoyl-ACP binds first (76, 77).

LpxH and LpxG

LpxH is a hydrolase that cleaves the pyrophosphate of UDP-2,3-bis(3R-hydroxymyristoyl)-glucosamine (UDP-DAG) in the fourth step of the Raetz pathway (1, 57, 81). Around 70% of Gram-negative bacteria utilize LpxH while *Chlamydiae* utilize the distantly related LpxG, and the rest utilize the nonhomologous enzyme LpxI (82, 83). LpxH was one of the primary focuses of my thesis research; therefore, further discussion of LpxH is reserved for Chapter 2.

LpxG was identified as a gene from *Chlamydia trachomatis* that could complement a *ΔlpxH* strain of *E. coli* (83). Activity assays showed that, like LpxH, LpxG is a Mn^{2+}-dependent pyrophosphatase that catalyzes the nucleophilic attack of water on the α-phosphate of UDP-DAG (83, 84). Consistently, mutation of the predicted metal-coordinating residue D59 to Ala significantly decreased activity (83). Also like LpxH, homology suggests that LpxG is a member of the Calcineurin-like metal-dependent phosphoesterase family (83). However, LpxG is predicted to have an additional N-

terminal transmembrane helix, and the sequence identity with LpxH is very low (11%) (83). Furthermore, while sequence alignments suggest that LpxG has a similarly located insertion in the Calcineurin-like metal-dependent phosphoesterase fold, the length and sequence of the insertions in LpxH and LpxG are different, suggesting that these proteins either diverged early in their evolution or arose by convergent evolution from different phylogenetic clades within the same enzyme family (85). The structure of LpxG has not been determined.

Another gene called *lpxH2* because of its sequence homology to *lpxH* is present in some Gram-negative bacteria without *lpxH* but can also be present in addition to *lpxH* (7, 86). Expression of LpxH2 could not complement the a *ΔlpxH* strain of *E. coli*, and the sequence of the substrate-binding helical cap domain of LpxH is not conserved in LpxH2 (85, 86). Thus, the function of *lpxH2* and its relevance to lipid A synthesis remains unknown.

LpxI

LpxI is a non-homologous alternative to LpxH present in some Gram-negative bacteria that lack LpxH and LpxG, particularly α-proteobacteria (7, 82). *LpxI* was discovered as a gene of unknown function present in the same operon as *lpxA*, *lpxD*, and *lpxB* in bacteria lacking *lpxH* (82). Unlike LpxH, LpxI does not utilize a catalytic metal center; though, it did require divalent cations, specifically Mg^{2+}, Mn^{2+}, or Co^{2+}, for optimal activity (82, 87). Furthermore, ^{18}O incorporation from $H_2{}^{18}O$ indicated that LpxI catalyzes the nucleophilic attack of water on the β-phosphate of UDP-DAG (82).

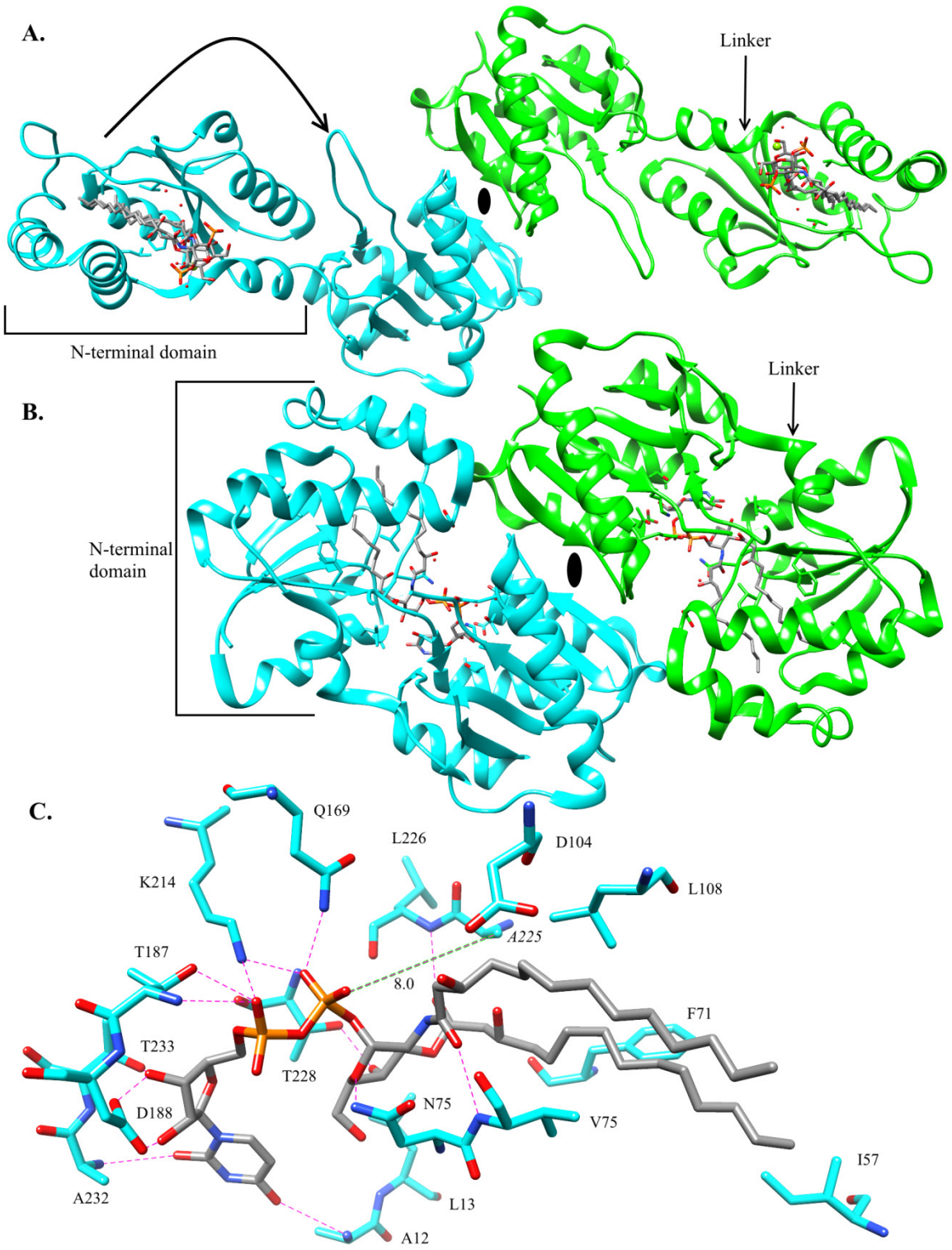


Figure 1.7: LpxI. A. LpxI dimer bound to lipid X (PDB: 4GGM) (87). **B.** LpxI-D225A dimer bound to UDP-DAG (PDB: 4J6E) (87). **C.** Residues involved in UDP-DAG binding to LpxI-D225A from **B.** Hydrogen-bond and salt-bridge interactions are show, and the distance between A225 and the β-phosphate

is highlighted. Distances in Ångstroms are as follows: 3.3 from 3'-acyl carbonyl to L226 N, 2.5 from ribose 2'-hydroxyl to D188 Oδ2, 2.3 from ribose 3'-hydrxyl to D188 Oδ1 and 3.0 to T233 Oγ, 3.3 from α-phosphate to T187 N and 2.8 to T187 Oγ and 3.3 to K214 Nζ, 2.6 from β-phosphate to K214 Nζ and 2.9 to Q169 Nε, 2.5 from glucosamine 4'-hydroxyl to T228 Oγ, 3.3 from uracil O2 to A232 N, 2.9 from glucosamine ring O to N74 Nδ, 3.4 from uracil O4 to A12 N, and 3.1 from 2'-acyl carbonyl to V75 N.

Crystal structures of *Caulobacter crescentus* LpxI and LpxI-D225A bound respectively to 2,3-bis(3R-hydroxymyristoyl)-glucosaminyl-1-phosphate (lipid X) and UDP-DAG revealed that LpxI has a novel two-domain fold wherein the N-terminal domain forms the binding pocket for the majority of the lipid substrate and the C-terminal domain forms the pyrophosphatase active site (87). The N-terminal domain is composed of a parallel β-sheet surrounded by 4 α-helices, and the C-terminal domain is composed of a six-stranded β-sheet surrounded by 4 α-helices and a small two-stranded β-sheet (87). The two domains are connected by a flexible linker that undergoes secondary structure rearrangement to allow the domains to come together and bring the substrate into the active site (Figure 1.7) (87). In the crystal structures, LpxI forms a dimer with two-fold symmetry (Figure 1.7), and LpxI also appeared to form a dimer in solution (87). However, the functional relevance of dimerization is unclear (87).

In the structure of LpxI-D225A in complex with UDP-DAG, uracil is bound by the backbone amides of residues 12 and 232, and ribose is hydrogen-bonded to D188 and T233 (87). The pyrophosphate forms an ion pair with K214 and hydrogen-bonds with Q169 and T187 (Figure 1.7) (87). Furthermore, the importance of T187 was supported by activity assays that showed T187A had 175-fold less activity than wild-type LpxI (87). Glucosamine is hydrogen-bonded to N74 and T288, and the carbonyls of the acyl chains are hydrogen-bonded to the backbone amides of residues 75 and 226 while the hydrocarbon chains bind in the complementary hydrophobic pocket of the N-terminal

domain (87). Although LpxI activity was stimulated by divalent cations, none are bound to the pyrophosphatase domain of either crystal structure nor are putative nucleophilic waters apparent (82, 87). Therefore, the exact mechanism of catalysis remains unclear. However, positioning of A225 suggests that D225 may position the nucleophilic water for attack on the β-phosphate and/or act as the catalytic base; in which case, the divalent cation may simply coordinate the unengaged side of the pyrophosphate to balance the negative charge to allow nucleophilic attack (87).

LpxB

LpxB is the glycosyltransferase that forms the base glucosamine disaccharide of lipid A, catalyzing the formation of the glycosidic bond between the anomeric carbon of the UDP-DAG glucosamine and the 6-hydroxyl of lipid X to form lipid A disaccharide (40). LpxB was one of the primary focuses of my thesis research; therefore, further discussion of LpxB is reserved for Chapter 3.

LpxK

LpxK is the kinase that phosphorylates the 4'-hydroxyl of the distal glucosamine of lipid A disaccharide to form lipid IV_A (88, 89). LpxK has been proposed to be the second regulatory point of the Raetz pathway: LpxK may be stimulated by unsaturated fatty acids, and its lipid A disaccharide substrate may stimulate the proteolysis of LpxC by FtsH (69). This regulation allows Gram-negative bacteria to balance phospholipid and LPS synthesis as both pathways compete for β-hydroxymyristoyl-ACP (68). High flux through the Raetz pathway depletes β-hydroxymyristoyl-ACP, limiting the production of unsaturated fatty acids, and decreased unsaturated fatty acid production leads to decreased LpxK activity, which causes a buildup of lipid A disaccharide (68, 69).

Buildup of this intermediate then leads to the proteolysis of LpxC, decreasing consumption of β-hydroxymyristoyl-ACP by the Raetz pathway (68, 69). In addition, LpxK is the final absolutely required step of the Raetz pathway as *E. coli* could transport lipid IV_A to the outer membrane and remain viable (5).

LpxK was most active with ATP and an equimolar concentration of divalent cation, showing little activity when divalent cations were removed with EDTA (88, 90). For the lipid substrate, LpxK was specific for the glucosamine disaccharide head-group, showing no activity for lipid X, UDP-DAG, or substrates already containing Kdo core oligosaccharide sugars (89). However, LpxK had little specificity for the number of acyl chains attached to the glucosamine disaccharide: LpxK could phosphorylate substrates with 2-6 acyl chains and substrates wherein the ester-linked chains were replaced with amide-linked chains (88, 89).

Apo, substrate analogue-bound, and product-bound crystal structures of *Aquifex aeolicus* LpxK have been determined (90–92). These structures reveal that LpxK is composed of two Rossmann-like domains with β-sheets surrounded by α-helices, and these domains are connected by two β-strands that are coextensive with the β-sheets of both domains (Figure 1.8) (91). The larger N-terminal domain contains the P-loop/Walker A and Walker B motifs thus binding ATP, and the C-terminal domain binds the lipid substrate (90–92). A hinge-like movement of the domains closes the distance between the binding pockets observed in the apo structure to bring the substrates into position for phosphoryl transfer (91).

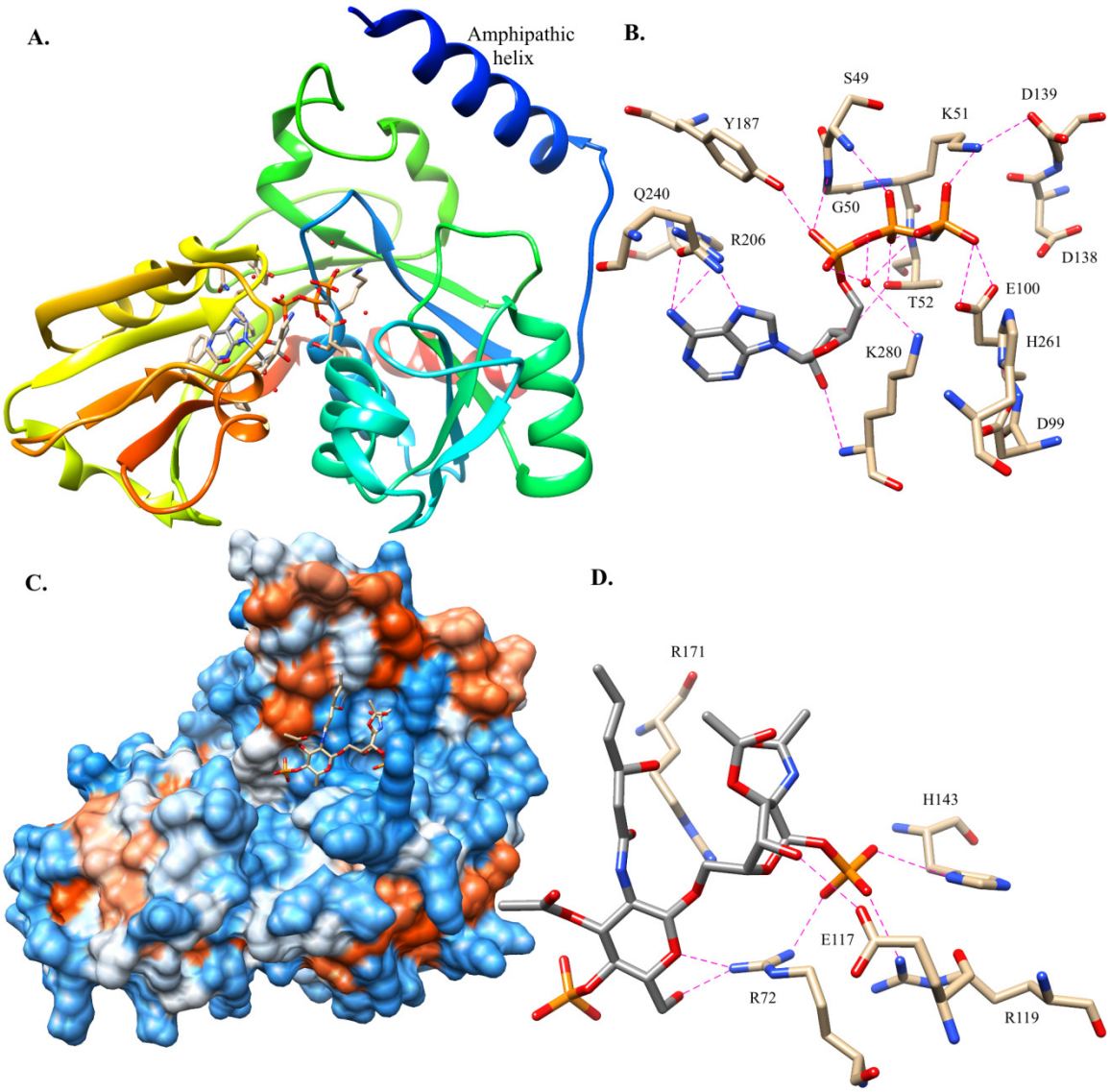


Figure 1.8: LpxK. A. LpxK bound to AMP-PCP ATP analogue (PDB: 4ITL) (90). Spectrum coloring begins with blue at the N-terminus. **B.** Residues involved in AMP-PCP binding from **A.** Distances for hydrogen-bond and salt-bridge interactions are shown with distances in Ångstroms as follows: 3.0 from adenine N6 to Q240 Oε and 3.5 to R206 Nη1, 3.0 from adenine N7 to Q240 Nε, 2.8 from ribose 3'-hydroxyl to T52 Oγ, 2.9 from β-phosphate to S49 N and 3.1 to K51 N and 2.7 to T52 N and 2.6 to T52 Oγ, 3.0 from α-phosphate to G50 N and 2.6 to Y187 Oη, 3.4 from ribose 2'-hydroxyl to K280 N, 2.8 from D99 Oδ1 to H261 Nδ, 3.0 from K51 Nζ to D139 Oδ1 and 2.7 to γ-phosphate, 2.9 and 2.8 from γ-phosphate to E100 Oε2/1, 3.2 from K280 Nζ to water, and 2.2, 3.0, and 3.0 from this water to the phosphates. **C.** Hydrophobicity surface of LpxK bound to lipid IV_A (PDB: 4LKV) (92). Hydrophobicity is shown in an orange-blue scale with blue representing the hydrophilic end. **D.** Residues involved in binding lipid IV_A from **C.** Distances in Ångstroms for hydrogen-bonds and salt-bridges are as follows: 2.4 from 6'-hydroxyl

to R72 Nη2 and 3.2 from the distal glucosamine ring O, 3.4 from 1-phosphate to R119 Nη1 and 3.2 to H143 Nδ and 2.9 from R72 Nη1, and 3.3 from 4-hydroxyl to E117 Oε.

In addition, LpxK has an N-terminal amphipathic helix and adjacent basic patch, which likely mediate membrane association; consistent with the presence of these structural elements, LpxK was optimally active at the critical micelle concentration of Triton X-100, and surface dilution of LpxK activity was observed at higher concentrations (90, 91). Moreover, the amphipathic helix was required for activity (92). LpxK was the first membrane active kinase of the P-loop-containing NTPase superfamily to be identified (91).

The AMP-PCP-bound structure of *A. aeolicus* LpxK shows that adenine is recognized by hydrogen-bonding to Q240 and R206 while ribose is bound by T52 and the backbone amide of K280 (90). Furthermore, the α- and β-phosphates are engaged in extensive hydrogen-bonding interactions with T52, Y187, the backbone amides of residues 49-52, and an ordered water bound to K280 (90). The γ-phosphate is positioned for transfer by a salt-bridge with K51 and hydrogen bonds to E100 and the ordered water (Figure 1.8) (90). Homology suggests that T52, E100, and D138 will coordinate Mg^{2+} in the complex with ATP-Mg (90, 91). The role of E100 in Mg^{2+} coordination rather than simply in hydrogen-bonding is supported by E100A/D/Q mutants wherein E100Q showed the greatest decrease in k_{cat} and E100D showed the least as Asp can also coordinate Mg^{2+}, and Ala may allow binding of a compensatory water (90). In addition, H261 is positioned to accept a proton from the attacking 4'-hydroxyl of lipid A disaccharide, and D99, which is hydrogen-bonded to H261, likely serves to increase its pK_a (90). Consistent with this role, D99A/E/N mutants showed the same trend as the E100 mutants (90). Kinetic experiments also confirmed the importance of K51, T52, and

H261, showing 3000-fold decreases in k_{cat} for K51A and T52A and an 850-fold decrease for H261A, and the Walker B motif residues D138 and D139 were observed to be critical for catalysis and ATP-Mg binding, decreasing k_{cat} 4700- and 8100-fold and increasing K_M 3.3- and 4-fold, respectively (90).

The structure of LpxK bound to its lipid IV_A product shows that LpxK binds the lipid IV_A glucosamines with the acyl chains extending into solvent and largely disordered, which explains the minimal specificity for the extent of acylation (88, 89, 92). The lipidic substrate binding site is on the same face as the N-terminal amphipathic helix, which suggests that LpxK may only partially extract lipid A disaccharide from the membrane during catalysis (92). Even with glucosamine disaccharide, there are relatively few interactions (92). The 1-phosphate is recognized by R72, R119, and H143; however, the proximal glucosamine is only hydrogen-bonded to E117, and the distal glucosamine is only hydrogen-bonded to R72 (Figure 1.8) (92). Kinetic experiments confirmed the importance for lipid A disaccharide binding of R72 and H143, as Ala mutants increased K_M 4- and 7-fold, respectively, but R119 was largely dispensable, showing no change in K_M and only a 67-fold decrease in k_{cat} (92). Finally, R171A increased K_M 20-fold, but R171 was not directly bound to lipid IV_A, which suggests that the lipid IV_A-bound structure may not fully recapitulate productive lipid A disaccharide binding (92).

WaaA (KdtA)

WaaA, also called KdtA, is the glycosyltransferase that, in *E. coli*, transfers two Kdo residues from CMP-Kdo: first to the 6'-hydroxyl of the distal glucosamine and then to the 4'-hydroxyl of the first Kdo (52, 53). While *E. coli* WaaA is thus a bifunctional glycosyltransferase, WaaA variants from different species transfer 1-4 Kdo residues (1,

93, 94). Activity assays also showed that, like LpxK, WaaA had little specificity for the extent of acylation, accepting substrates with 3-6 acyl chains, nor was the 1-phosphate required (52, 53, 94). However, the 4'-phosphate was important as lipid A disaccharide was a poor substrate for *E. coli* WaaA though even lipid A disaccharide could be glycosylated to a lesser extent by *Aquifex aeolicus* WaaA (52, 53, 94). Conversely, WaaA was highly specific for the donor substrate, showing activity and specific binding for CMP but no other nucleoside monophosphates (53, 94). Interestingly, WaaA is another point of regulation in the Raetz pathway as WaaA was also degraded by the FtsH protease that targets LpxC (95).

Crystal structures of the monofunctional WaaA from *A. aeolicus* revealed that WaaA has a typical glycosyltransferase B superfamily (GT-B) fold with two Rossmann-like domains (Figure 1.9) (96). Moreover, the N-terminal domain of WaaA has a basic and hydrophobic surface as observed for other membrane surface-active GT-B enzymes (including LpxB as discussed in Chapter 3), which is thought to be required for productive membrane association (96–100). As expected for a GT-B enzyme, CMP was observed to bind to the C-terminal domain, and a putative lipid IV_A binding pocket was identified on the N-terminal domain (96, 98). WaaA appears to have been crystallized in an open conformation with a wide cleft separating the domains; therefore, a hinge-like movement of the domains may be required to bring the substrates into position for nucleophilic attack (96–99, 101, 102).

The CMP-bound structure of *A. aeolicus* WaaA shows that cytosine binds in a hydrophobic pocket formed by F247 and L250 and is recognized by hydrogen-bonds with the backbone carbonyl of P211 and amide of G248 (96).

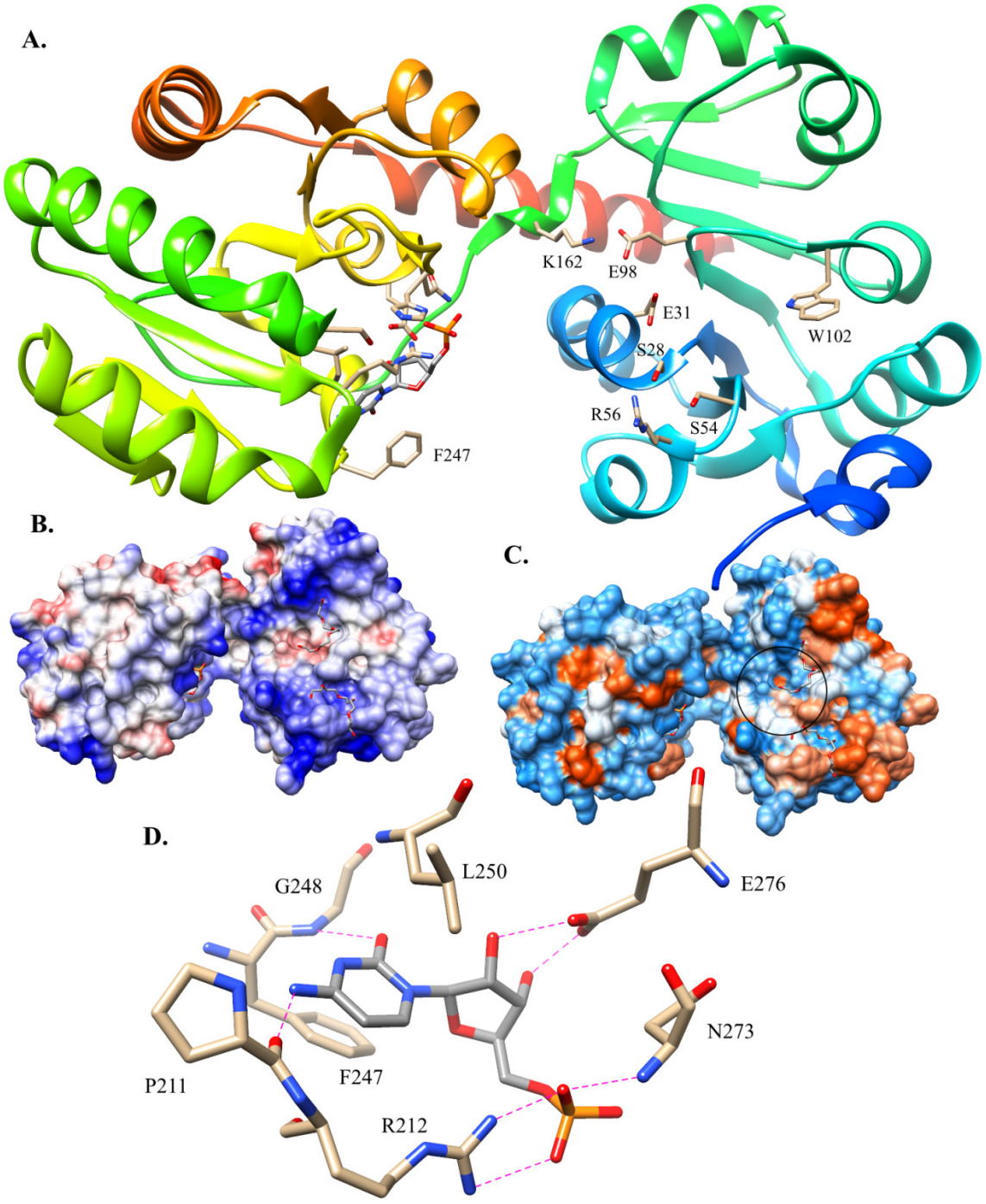


Figure 1.9: WaaA. A. WaaA bound to CMP product (PDB: 2XCU) (96). Spectrum coloring ends with red at the C-terminus. **B.** Columbic surface of WaaA from **A.** Charge is shown with a blue-red scale with blue at positive end. Bound polyethylene glycol molecules are shown. **C.** Hydrophobicity surface of WaaA from **A.** Hydrophobicity is shown on an orange-blue scale with orange as the hydrophobic end. The putative lipid IV_A-binding pocket is circled. **D.** Residues involved in CMP binding from **A.** Hydrogen-bond and

salt-bridge interactions are shown with the following distances in Ångstroms: 3.0 from phosphate to N273 N and 2.6 and 3.2 to R212 Nη1/2, 2.5 from E276 Oε1 to ribose 3'-hydroxyl and 2.8 to the 2'-hydroxyl, 2.9 from cytosine O2 to G248 N, and 3.1 from cytosine N4 to P211 O.

Ribose is hydrogen-bonded to E276, and the phosphate is engaged by R212 and N273 (Figure 1.9) (96). Activity assays confirmed the importance of R212, which may stabilize the negative charge of the CMP leaving group during catalysis (96). In addition, activity assays of Ala mutants confirmed the importance of residues G30, E31, E98, and K162 in the putative lipid IV_A binding pocket (96). E31 appears well positioned to act as the catalytic base by accepting a proton from the 6'-hydroxyl of the distal glucosamine, and G30 could prevent steric clash from residue 30 that could interfere with proper alignment for nucleophilic attack (96). If this role is correct for E31, then E98 and K162 may be important for positioning Kdo for transfer, possibly by contacting its 1'-carboxlyate, and residues S28, S54, and R56 may bind the 1-phosphate of lipid IV_A (96). Mutation of all 3 of these putative 1-phosphate binding residues to Ala significantly reduced activity (96). Finally, increase in Trp-fluorescence upon lipid IV_A binding in wild-type but not W102A WaaA suggested that the acyl chains bind on top of W102, shielding it from solvent quenching (96).

LpxL and LpxP

In *E. coli*, LpxL is the acyltransferase that adds a lauroyl group to the β-hydroxyl of the 2'-β-hydroxymyristate on the distal glucosamine of Kdo_2-lipid IV_A (103, 104). Enzymatic assays showed that LpxL was a membrane surface active enzyme, and a predicted N-terminal transmembrane helix was required for activity *in vivo* (104). In addition, enzymatic assays showed that LpxL was selective for the presence of the Kdo residues, as its activity for lipid IV_A was 6000-fold lower (104). Furthermore, LpxL was

selective for lauroyl-ACP: LpxL was 5% as active with lauroyl-coenzyme A (CoA), showed further reduced activity with decanoyl-CoA and myristoyl-CoA, and incorporated little or no acyl chains 16 carbons or longer from acyl-CoA (104).

Conservation of catalytic residues suggested that LpxL is related to the glycerol-3-phosphate acyltransferase (GPAT) family (104). Consistent with this hypothesis, mutation to Ala of H132 and E137 of the predicted catalytic dyad in *E. coli* LpxL decreased activity 1000- and 3000-fold, respectively (104). Moreover, Ala mutations of conserved R169 and D200 residues decreased activity 170- and 15-fold (104). However, P238A had little effect even though this Pro is important in the GPAT family, and this Pro was not conserved in the related enzyme LpxP (104).

E. coli also have a cold-shock inducible alternative to *lpxL* (*lpxP*) that replaces LpxL when cells are grown at 12^{o}C (105). Like LpxL, LpxP was specific for the presence of the Kdo groups of Kdo_2-lipid IV_A as LpxP failed to acylate lipid IV_A (105). However, LpxP transfers a palmitoleoyl chain from ACP instead of a lauroyl chain (105). In addition, LpxP was highly specific for (16:1) palmitoleoyl-ACP, showing little activity with shorter or saturated acyl chains or with palmitoleoyl-CoA (105). While LpxL could substitute for LpxP at 12^{o}C in a *ΔlpxP E. coli* strain, this strain showed increased sensitivity to rifampicin and vancomycin at 12^{o}C, which suggests that incorporation of unsaturated acyl chains into the outer leaflet of the outer membrane by LpxP maintains the selective barrier at low temperature (106). Neither LpxL nor LpxP have been structurally characterized, but the related enzyme LpxM has been (107).

LpxM

In *E. coli*, LpxM is the acyltransferase that attaches a myristoyl group to the β-hydroxyl of the 3'-β-hydroxymyristate on the distal glucosamine of the LpxL product (Kdo_2-lipid V_A) to produce the end product of the Raetz pathway, hexa-acylated Kdo_2-lipid A (108). Enzymatic assays suggested that *E. coli* LpxM was specific for the penta-acylated substrate as LpxM acylated Kdo_2-lipid IV_A and lipid IV_A much more slowly (108). However, while LpxM showed poor activity for decanoyl-ACP, β-hydroxymyristoyl-ACP, palmitoyl-ACP, and palmitoleoyl-ACP, LpxM appeared to only have a slight preference for myristoyl-ACP over lauroyl-ACP (108).

The crystal structure of LpxM from *Acinetobacter baumannii* reveals that LpxM has a unique fold with a twisted, 7-stranded β-sheet surrounded by 10 α-helices that form a deep hydrophobic pocket and with an N-terminal transmembrane/membrane-association helix (107). Unlike *E. coli* LpxM, *A. baumannii* LpxM transfers lauroyl chains to the β-hydroxyls of the acyl chains at the 2- and 3'-positions of the glucosamine disaccharide; therefore, *A. baumannii* LpxM is a bifunctional acyltransferase that transfers secondary lauroyl groups to both glucosamines to produce hepta-acylated lipid A (107). Kinetic experiments showed substrate inhibition by lauroyl-ACP, suggesting that LpxM utilizes an ordered sequential mechanism in which the acyl acceptor must bind first (107). Consistent with the bifunctionality of *A. baumannii* LpxM, substrate inhibition was best fit by a two-site binding model wherein only one lauroyl-ACP is inhibitory, and substrate inhibition was alleviated in the K282E/K285E double mutant of a putative ACP-binding site at the top of the lipid-binding pocket (107).

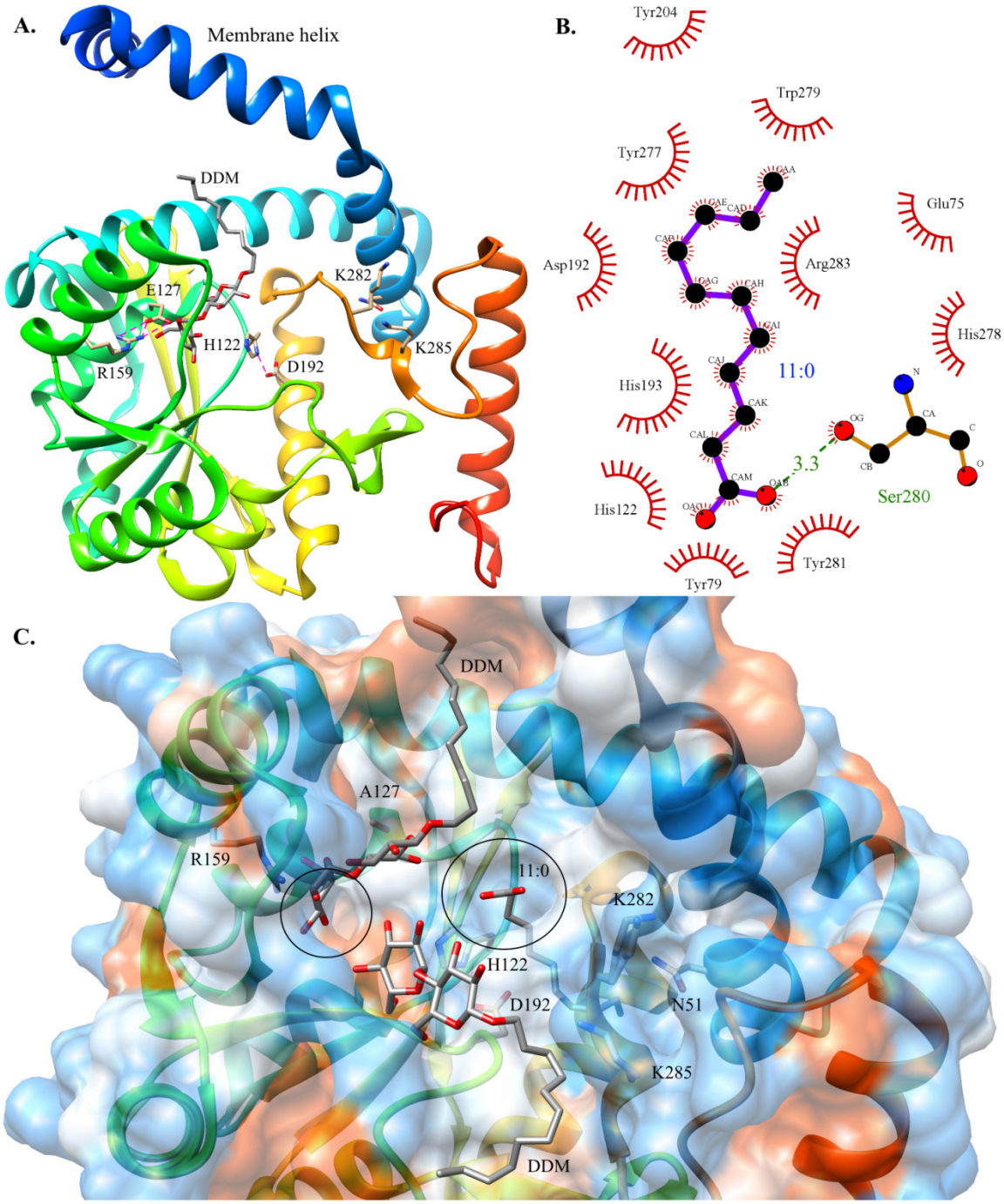


Figure 1.10: LpxM. A. Ribbon diagram of LpxM (PDB: 5KN7) (107). Spectrum coloring ends with red at the C-terminus. Interactions between putative catalytic residues are shown. Distances in Ångstroms are 2.6 from H122 Nδ to D192 Oδ1, 2.4 from E127 Oε1 to R159 Nη2 and 3.4 to Nη1, and 3.2 from E127 Oε2 to R159 Nη2. **B.** Ligplot+ diagram of fatty acid binding to LpxM (PDB: 5KNK) with distance shown in Ångstroms (78, 107). Hydrophobicity surface of LpxM (PDB: 5KNK). Hydrophobicity is shown on an

orange-blue scale with blue as the hydrophilic end. The entrances to putative hydrocarbon-ruler channels are circled.

Like LpxL, LpxM has a conserved His and Glu (H122 and E127 in *A. baumannii*) that, in distantly related acyltransferases, form the catalytic dyad with the His acting as the catalytic base that accepts a proton from the nucleophilic hydroxyl (107, 109). Kinetic experiments confirmed that E127 was critical for activity (107). In addition, LpxM has a conserved Arg and Asp (R159 and D192); R159 was also confirmed to be critical for activity (107). The crystal structure of *A. baumannii* LpxM reveals the location of these residues in the hydrophobic pocket, and the positions of these residues suggest that, at least in this bifunctional LpxM, the roles of the conserved HX_4E/D motif have been separated to form two catalytic dyads that catalyze acyl transfer to the 2- and 3'-β-hydroxymyristoyl chains (Figure 1.10) (107). In the glycerol-3-phosphate acyltransferase from *Cucurbita moschata*, the His and Asp of the HX_4E/D catalytic dyad are hydrogen-bonded, suggesting that Asp increases the pK_a of His (110). However, in *A. baumannii* LpxM, H122 and E127 are separated by 7.6 Å, and H122 is hydrogen-bonded to D192 (107). Moreover, R159 forms an ion pair with E127 (107). Therefore, H122 and D192 may serve as the catalytic dyad for the transfer of one lauroyl group, and E127 and R159 may act form the catalytic dyad for the transfer of the second lauroyl group. Finally, the structure of the *A. baumannii* LpxM E127A mutant was resolved with an acyl chain (modeled as undecanoic acid) bound in one of two particularly deep, narrow channels in the active site (Figure 1.10) (107). These channels may act as hydrocarbon rulers to control the length of the transferred acyl chains similar to those in LpxA and LpxD as discussed above (107). Consistent with the hypothesis that H122/D192 and R159/E127 form two catalytic dyads for separate acyl transfer reactions, one of these putative

hydrocarbon-ruler channels is near H122 and D192, and the other channel is near R159 and E127 (Figure 1.10).

LpxJ and LpxN

Some bacteria encoding *lpxL* but not *lpxM* instead encode a distant homologue of *lpxL*, *lpxM*, and *lpxP* designated *lpxJ* (111). LpxJ enzymes from three ε-proteobacteria species, *Helicobacter pylori*, *Campylobacter jejuni*, and *Wolinella succinogenes*, were characterized, showing a range of substrate specificities (111). Ultraviolet photon disassociation mass spectrometry combined with activity assays confirmed that *H. pylori* LpxJ acylated the β-hydroxyl of the 3'-acyl chain, like *E. coli* LpxM, and not the acyl chains of the proximal glucosamine (111). However, none of these enzymes required the presence of the Kdo groups, and *H. pylori* LpxJ was not specific for the presence of the 2'-secondary acyl chain added by LpxL while the other LpxJ variants were only active on Kdo_2-lipid IV_A or lipid IV_A, not the LpxL product (111). *H. pylori* and *W. succinogenes* LpxJ were most active with lauroyl-ACP, but *H. pylori* LpxJ also utilized myristoyl-ACP (111). On the other hand, *C. jejuni* LpxJ was most active with palmitoyl- or stearoyl-ACP (111). In addition, *Vibrio cholerae* utilizes another alternative to LpxM, designated LpxN, that transfers a β-hydroxylauroyl chain to the β-hydroxyl of the 3'-acyl chain (36). Thus, the different activities and substrate specificities of the *lpxL/M* homologues encoded by different bacterial species result in diversity in the length, extent, and position of acylation in the final lipid A product of the Raetz pathway in different species (1).

1.4 LPS Synthesis and Transport

Core Oligosaccharide Synthesis

Following the completion of the Raetz pathway, the core oligosaccharide is completed at the cytoplasmic side of the inner membrane (1, 112). The core oligosaccharide can be divided into inner and outer cores, and the inner core tends to be conserved within a species or genus while species typically have a small number of different outer cores synthesized in different strains (1, 112). As described above, the first two Kdo residues are added during the Raetz pathway, and these sugar residues are the most conserved moiety of the inner core with all known core oligosaccharides containing Kdo (1, 52, 53). In addition, the inner core typically contains L-glycero-D-mannoheptose (Hep) or, less frequently, the D-glycero-D-mannoheptose isomer (1). In *E. coli*, the first and second Hep residues (Figure 1.11) are added sequentially to the first Kdo by WaaC and WaaF, respectively, and a branching Hep is added to the second Hep by WaaQ (1, 113). The first and second Hep residues are phosphorylated by WaaP and WaaY, respectively, and the activity of these kinases is dependent on the addition of the first outer core residue by WaaG (1, 114). These phosphorylated inner core residues are thought to be important for crosslinking of LPS molecules by divalent cations, and loss of phosphorylation increases sensitivity of *E. coli* to detergent and some antibiotics (1, 114). Moreover, the second Hep is the attachment site of this first outer core residue, which is glucose in *E. coli* (1, 115). In *E. coli* K-12 the first glucose is the site of attachment of a branching galactose by WaaB and a second glucose residue by WaaO (115). This WaaO product is a branch point in core synthesis in *E. coli* K-12: either WaaR adds a third glucose to the second glucose, or WaaZ adds a third Kdo to the second Kdo (11).

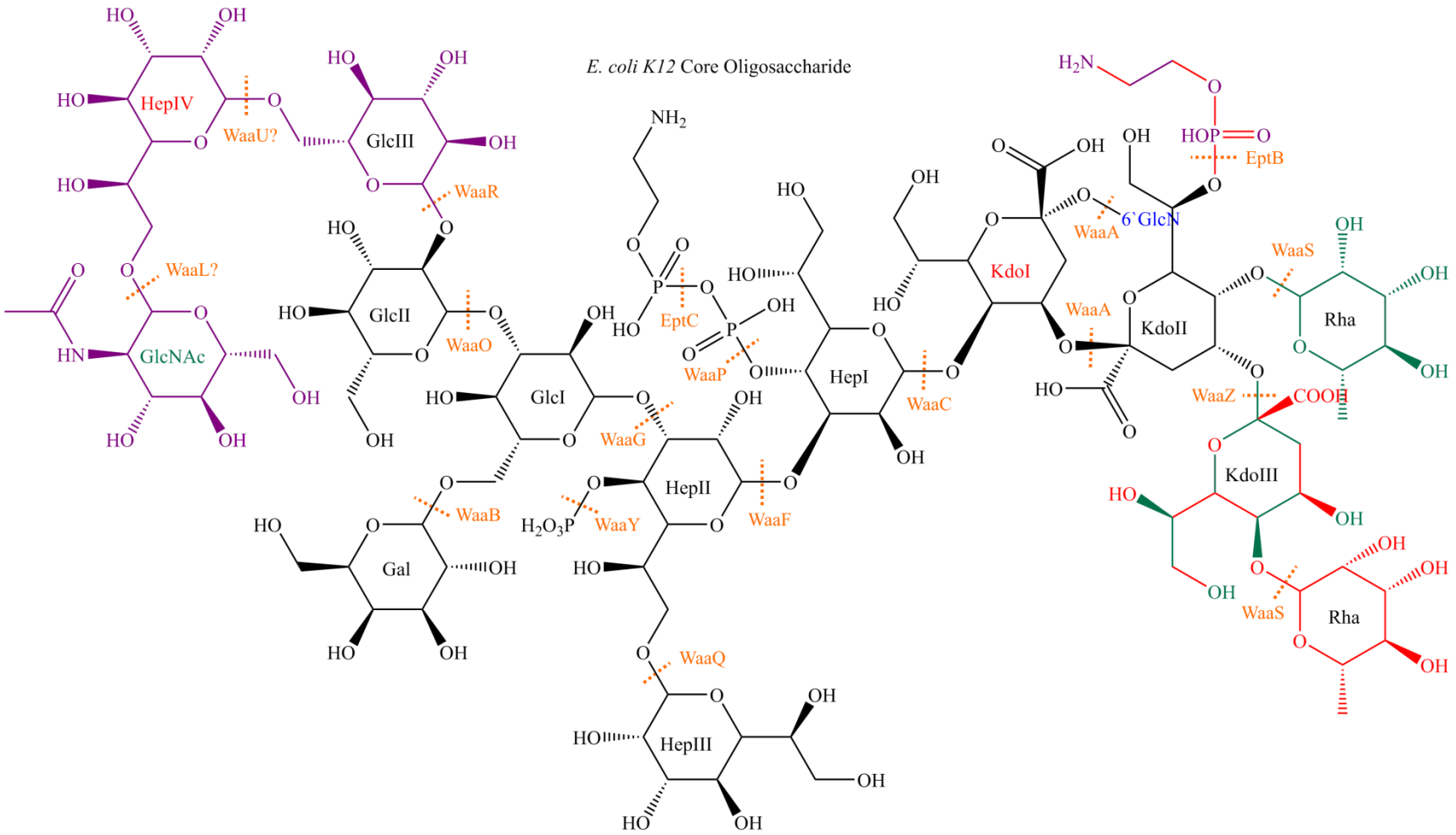


Figure 1.11: Core oligosaccharide synthesis. Coloring of core oligosaccharide is the same as in Figure 1.1. Transferases responsible for the formation of bonds linking residues in the K-12 core oligosaccharide are marked in orange (1, 11, 12, 113–118).

Competition between these pathways appears to be controlled by the relative levels of WaaR and WaaZ, which are dependent on growth conditions with phosphate limitation favoring WaaZ and optimal growth conditions favoring WaaR (11). Following addition of the third Kdo, WaaS can attach L-rhamnose (Rha) to the second Kdo or, if the second Kdo has already been modified with phosphoethanolamine by EptB, to the third Kdo (11). In the WaaR pathway, a fourth heptose is attached to the third glucose to complete the outer core (Figure 1.11) (1). Sequence similarity with other heptosyltransferases suggests that this reaction is catalyzed by WaaU (117).

Transport of Core-Lipid A across the Inner Membrane by MsbA

Following the completion of the core oligosaccharide, the core-lipid A molecule is flipped across the inner membrane by the ATP-binding cassette (ABC) transporter MsbA (5, 112, 119). Decreased activity of MsbA for earlier intermediates of LPS synthesis lacking core sugars or secondary acyl chains likely helps to prevent premature transport (5). Structures of MsbA obtained by cryogenic electron microscopy (cryo-EM) and x-ray crystallography have provided insight into the mechanism of transport (120, 121). These structures show that MsbA forms a functional dimer wherein the transmembrane domains are comprised of 4 transmembrane helices from one subunit and 2 transmembrane helices (TM4 and TM5) from the opposite subunit (Figure 1.12) (120, 121). A cytoplasmic P-loop containing ATPase domain is appended to TM6, and cytoplasmic helices between TM2 and TM3 and between TM4 and TM5 contact the ATPase domains (120, 121). The helix between TM2 and TM3 contacts its own ATPase domain at the ATP-binding site, and the helix between TM4 and TM5 contacts the ATPase domain of the opposite subunit (Figure 1.12) (120, 121).

The ATPase domains also act as a functional dimer: the 3.7 Å crystal structure of *Salmonella* MsbA bound to a non-hydrolyzable ATP analogue (AMP-PNP) shows that ATP binds between the ATPase domains (Figure 1.12C) (120). One half of the ATPase active site formed by the loops of two β-sheets and an α-helix sandwiched between them from one ATPase domain, and the other half is contributed by the α-helical side of the other ATPase domain (120). In this structure, the transmembrane domain assumes a periplasmic open conformation (Figure 1.12C) (120).

A 4.2 Å cryo-EM structure of *E. coli* MsbA in the cytoplasmic open conformation shows core-lipid A bound inside the transmembrane domains with partially resolved acyl chains and core oligosaccharide resolved out to the three inner core Hep residues (Figure 1.12) (121). The core-lipid A binding pocket is separated into hydrophobic lipid A-binding and hydrophilic oligosaccharide-binding regions by a ring of primarily basic residues that recognize the phosphorylated glucosamine disaccharide of lipid A (121). The acyl chains are directed toward the periplasm suggesting that suggesting that flipping occurs after a conformational change to open the periplasmic side of MsbA; however, examination of the hydrophobicity of MsbA (Figure 1.12B) suggests that the requisite translation has already occurred with the acyl chains already extending to the periplasmic leaflet of the inner membrane (121). A second 4.8 Å cryo-EM structure, determined in the presence of ADP and vanadate, shows MsbA in an occluded conformation with parallel transmembrane helices (121).

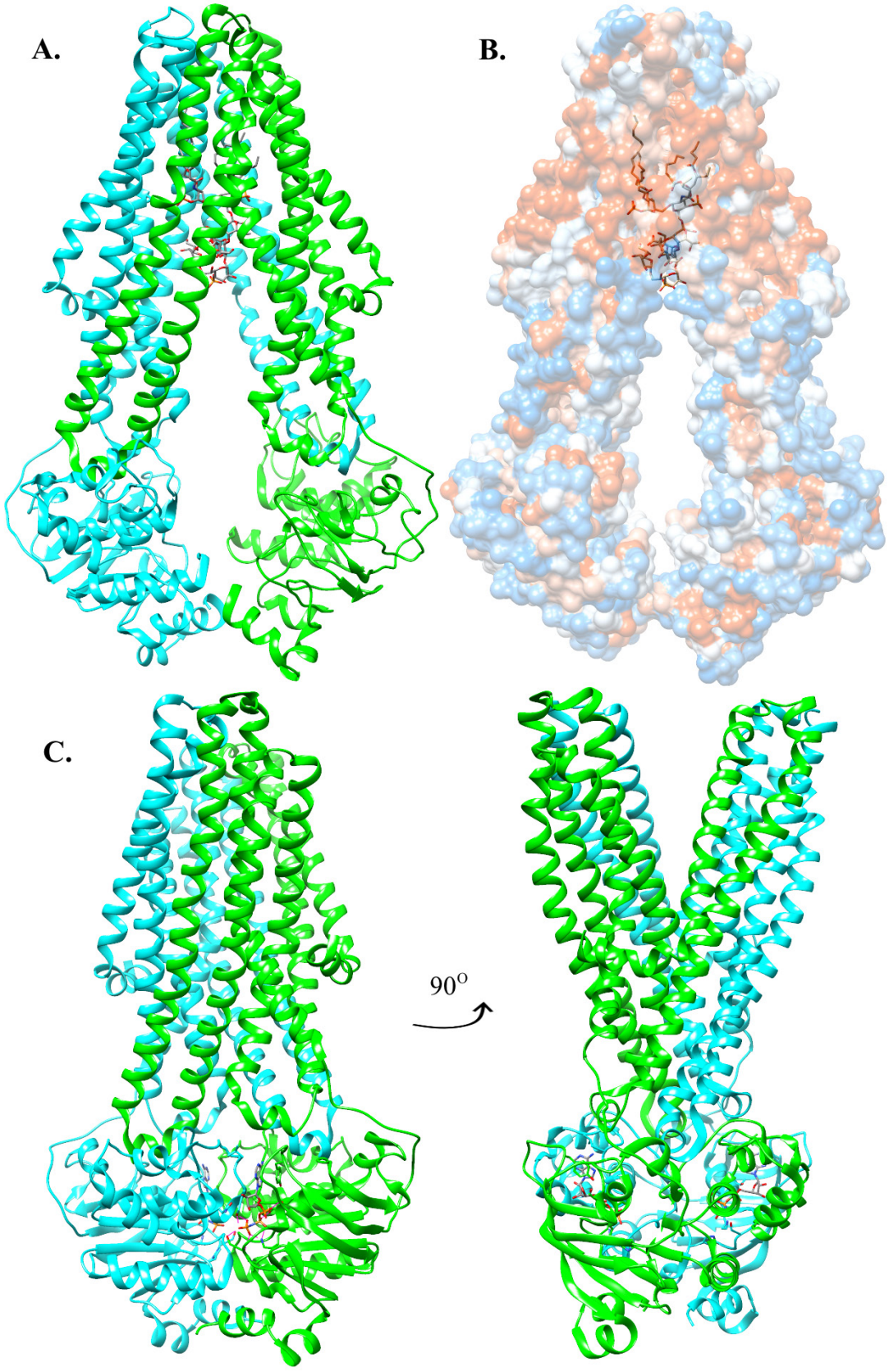

A.
B.
C.
90°

Figure 1.12: MsbA. A. Cytoplasmic-open conformation of MsbA bound to LPS (PDB: 5TV4) (121). **B.** Hydrophobicity surface of MsbA in **A.** Orange is the hydrophobic end of the scale, and blue is the hydrophilic end. **C.** Periplasmic-open conformation of MsbA bound to AMP-PNP ATP analogue (PDB: 3B60) (120).

Therefore, the structures of MsbA suggest a mechanism wherein core-lipid A binds the cytoplasmic open conformation, and ATP binding brings the ATPase domains together, which is coupled to a conformational change in the transmembrane domains from cytoplasmic open to periplasmic open (120, 121). Core-lipid A then flips and diffuses into the periplasmic leaflet (120, 121). Next, ATP hydrolysis is coupled to a periplasmic open to occluded conformational change, which prevents the transporter from acting in reverse (121). Finally, release of phosphate and/or ADP allows MsbA to return to the starting cytoplasmic open conformation (120, 121).

O-antigen Polysaccharide Synthesis and Transport

The O-antigen is a repeating polysaccharide that varies between bacterial strains (1). Different O-antigens are synthesized by one of three divergent pathways: Wzy dependent, ABC transporter dependent, and synthase dependent (1, 112). However, O-antigen synthesis has a few conserved features: O-antigen is built on undecaprenol phosphate (und-P), and sugar residues are transferred to this lipid from nucleotide-charged sugar donors at the cytoplasmic face of the inner membrane (1, 112). Furthermore, a sugar-1-phosphate [1-phospho-N-acetyl-glucosamine (P-GlcNAc) in *E. coli*] is transferred to und-P to form a phospho-anhydride product (GlcNAc-PP-und in *E. coli*) (1). Finally, the completed O-antigen is ligated to core-lipid A at the periplasmic face of the inner membrane to form completed LPS by WaaL, which is specific for core oligosaccharides but not O-antigens (1). The lack of polysaccharide specificity of WaaL

also allows it to attach other capsular polysaccharides (K-antigens and colonic acid in *E. coli*) to core-lipid A (118, 122). Moreover, while *E. coli* K-12 does not synthesize O-antigen, conservation of the machinery shared with other polysaccharide synthesis pathways still enables this strain to synthesize GlcNAc-PP-und and transport it across the inner membrane (117, 122, 123). Therefore, WaaL may also transfer N-acetyl-glucosamine to the terminal Hep of the K-12 outer core (Figure 1.11) (117, 118). Studies of the topology of WaaL indicated that the ligase contained 12 transmembrane segments and 2 major periplasmic segments, and enzymatic assays showed that a conserved Arg in a short periplasmic segment and a conserved His in the largest periplasmic segment were essential for activity (124, 125). Examination of the reactants and products of the WaaL reaction suggest that WaaL catalyzes an S_N2 substitution reaction wherein the terminal residue of the core oligosaccharide (7'-hydroxyl of the fourth Hep in *E. coli* K-12) attacks the anomeric carbon of the basal residue of the O-antigen with und-PP acting as the leaving group (125). The positions of these residues and the metal-independent activity of WaaL make it tempting to speculate that the conserved Arg and His could bind the pyrophosphate of O-antigen-PP-und and act as the catalytic base for the attacking hydroxyl, respectively (124, 125).

The Wzy dependent pathway is unique from the other pathways in that single repeat units are synthesized and transported across the inner membrane (1). Once the repeat unit has been synthesized by its specific glycosyltransferases, the repeat unit-PP-und is flipped across the inner membrane by a Wzx flippase (126). Little is known about the mechanism of Wzx flippases; even whether they require energy is a matter of contention though *in vitro* transport of a soluble analogue of GlcNAc-PP-und was not

affected by the addition of ATP, NADH, or ionophores (123, 126). However, Wzx flippases are thought to have 12 transmembrane segments, and they tend to be specific for a particular O-antigen repeat unit with the first sugar attached to und-PP playing the most important role in substrate specificity (124, 126). Next, the O-antigen repeat units are polymerized at the periplasmic face of the inner membrane on a single und-PP by Wzy, which catalyzes the transfer of the growing polysaccharide to the free end of the new repeat unit-PP-und (1). Topology studies of Wzy from *Pseudomonas aeruginosa* PAO1 suggested that the polymerase contained 14 transmembrane segments and 4 major periplasmic segments (124). Two of these periplasmic segments were found to contain $RX_{10}G$ motifs with highly conserved Arg residues that were essential for activity (127, 128).

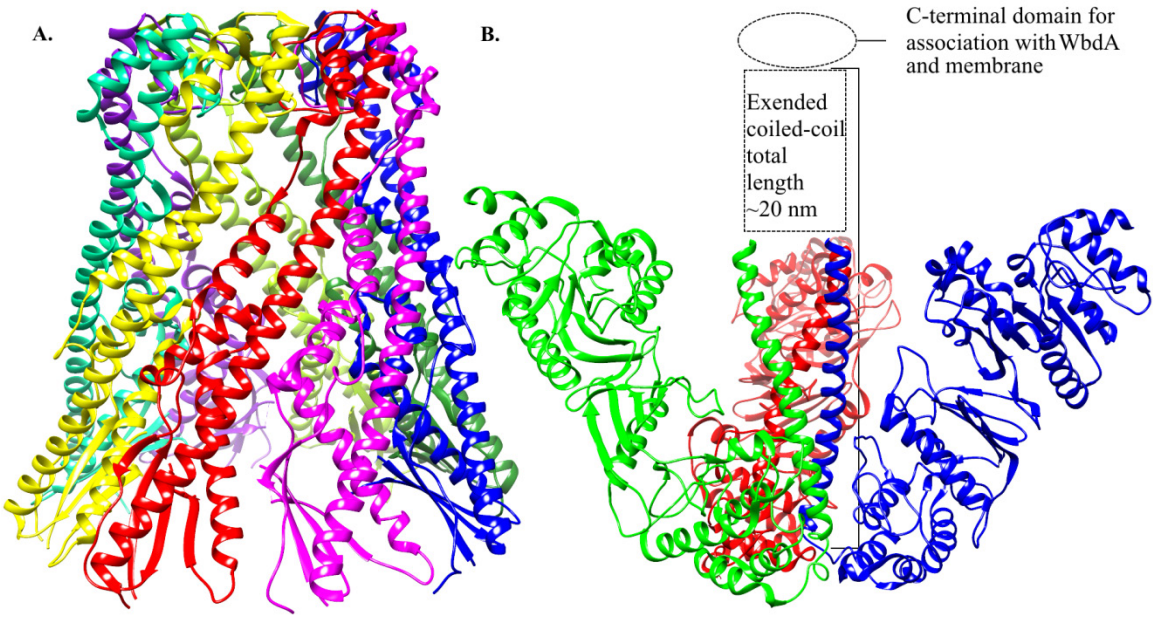


Figure 1.13: Proteins that control O-antigen modal length. A. Octamer of WzzB periplasmic domains (PDB:4E29) (129). **B.** WbdD trimer (PDB: 4UW0) with the location of the unresolved extended coiled-coil domain and uncharacterized C-terminal domain marked by a rectangle and oval, respectively (130, 131).

The presence of two similar motifs in Wzy corresponding with two similar substrates and the conservation of Arg suggests that the Arg residues could be involved in binding a common motif in Wzy substrates such as the pyrophosphate of oligo-PP-und (124, 127, 128). In addition to a specific sugar composition, O-antigens have a strain specific modal length, and this is controlled by Wzz (1, 129, 132). While crystal and cryo-EM structures of Wzz length regulators from multiple species have been determined, the mechanism of length regulation has remained obscure (129, 132, 133). Wzz is primarily a periplasmic protein with N- and C-terminal transmembrane helices, and the periplasmic region forms bell-shaped oligomers with 5-12 subunits though open trimers have also been observed (129, 132, 133). The periplasmic region has an α/β-domain that forms the base of Wzz adjacent to the membrane and an extended α-helical domain that forms a twisted helix bundle in oligomers (Figure 1.13) (129, 132, 133). *In vivo* assays of Wzz mutants and chimeras indicated that surface-exposed residues at the top of the α-helical domain controlled O-antigen length, but the role of oligomerization in length regulation (if any) has not been established (129, 132).

In the ABC transporter dependent pathway, the O-antigen is synthesized continuously on a single und-P carrier at the cytosolic face of the inner membrane (1, 134). Consequently, the initial 1-phospho-sugar and second adaptor residue are not repeated, forming the adaptor region before the repeat region of the O-antigen (1, 134). Moreover, ABC transporter dependent O-antigens can contain terminal modifications to the free end of the O-antigen that act as a form of quality control to ensure only completed O-antigen is transported across the inner membrane (1, 134, 135). In terminally modified O-antigens synthesized by the ABC transporter dependent pathway,

the strain specific modal length of the O-antigen is controlled by the enzyme that makes the terminal modifications, which halt the activity of the glycosyltransferases that add the repeat units (130, 134). Crystal structures and small angle x-ray scattering (SAXS) data of terminating enzyme of *E. coli* O9a (WbdD) revealed that the catalytic domain of WbdD is separated from a structurally uncharacterized C-terminal domain by an extended α-helical domain that forms a ~200 Å coiled-coil in WbdD trimers (Figure 1.13) (130). The C-terminal domain mediated association with the membrane and the glycosyltransferase that synthesizes the O9a repeat units (WbdA) (131). Deletions that decreased the length of the coiled-coil domain decreased the modal length of the O-antigen *in vivo*, and the opposite effect was observed for insertions designed to increase the length of the coiled-coil (130). Therefore, these data support a model wherein O-antigen is lengthened by WbdA until long enough to reach from the membrane to the catalytic domain of WbdD at which point termination depends on competition between WbdA and WbdD for the free end of O-antigen (130). However, some ABC transporter-dependent O-antigens lack terminal modifications, and in this case, the modal length is controlled by the stoichiometry of the glycosyltransferases that synthesize the repeat units and of the ABC transporter (134). Finally, the completed O-antigen is transported across the membrane by the ABC transporter (1, 134). The structure of the heterotetramer complex of the Wzm permease and the N-terminal ATPase domain of Wzt from *Aquifex aeolicus* and the structure of the C-terminal carbohydrate-binding domain of Wzt from *E. coli* O9a have been solved by x-ray crystallography (135, 136).

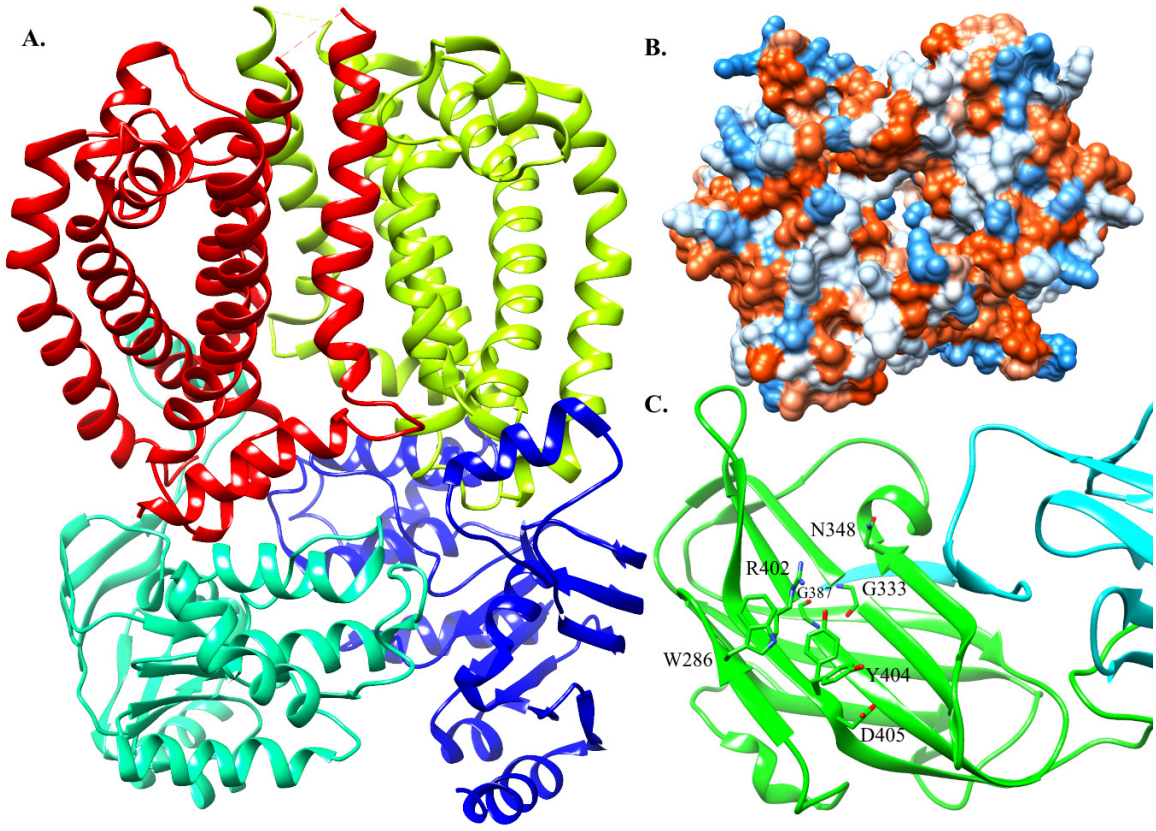


Figure 1.14: O-antigen ABC transporter. A. Wzm-Wzt heterotetramer ABC transporter from *Aquifex aeolicus* (PDB: 6AN7) (136). The Wzm subunits are shown in red and yellow, and the ATPase domain of the Wzt subunits are shown in cyan and blue. **B.** Hydrophobicity surface of Wzm-Wzt heterotetramer from **A.** showing periplasmic end of continuous channel. Hydrophobicity is shown on an orange-blue scale with orange as the hydrophobic end. The minimum radius of the channel is 3.5 Å (136). **C.** C-terminal carbohydrate-binding domain of Wzt from *E. coli* O9a (PDB: 2R5O) (135). The domain was crystallized as a C-terminally swapped dimer, but this C-terminal swap may not occur in the full Wzm-Wzt complex. Residues found to be involved in specific recognition of terminally modified O9a O-antigen are shown (135).

The Wzm-Wzt ABC transporter complex was solved in an open conformation and forms a continuous channel with a minimum radius of 3.5 Å, suggesting that the polysaccharide can occupy the channel through multiple rounds of ATP hydrolysis (Figure 1.14) (136). The Wzm channel is lined with several aromatic residues that may form stacking interactions with the sugar rings (136). The Wzt C-terminal domain forms an

immunoglobulin-like β-sandwich with a groove containing several residues that were shown to be involved in recognition of the terminally modified O9a polysaccharide *in vivo* and *in vitro* (Figure 1.14) (135).

As in the ABC transporter dependent pathway, synthase-dependent O-antigen is synthesized on a single und-P carrier and thus contains an unrepeated adaptor region (1). The synthase (WbbF) is thought to simultaneously add repeat units and transport the growing O-antigen across the membrane (1). However, only O:54 O-antigen of *Salmonella enterica* is known to be synthesized by this pathway, and it has thus not been studied extensively (1).

Transport of LPS to the Outer Membrane

Completed LPS is transported from the periplasmic leaflet of the inner membrane to the outer leaflet of the outer membrane by the Lpt complex, which is comprised of proteins LptA-G and spans from the cytoplasmic face of the inner membrane to the extracellular face of the outer membrane (137–140). This complex, which is formed by β-jellyroll domains in LptF, LptG, LptC, LptA, and LptD, forms a continuous hydrophobic slide for LPS across the periplasm, and transport is driven by the ATPase activity of LptB (139–148). Extensive structural and functional data have been obtained to support this transenvelope protein bridge mechanism.

Some of the earliest functional evidence for the transport of LPS via a transenvelope bridge was obtained by the study of lipid transport from inner membrane spheroplasts (149). These spheroplasts were found to contain pieces of outer membrane that could be separated by centrifugation (149). Newly synthesized (^{14}C-labeled) LPS was transported to the outer membrane in an MsbA-dependent, periplasmic extract-

independent manner, and this contrasted with the transport of lipoproteins, which could be released from spheroplasts by the addition of the soluble periplasmic carrier protein LolA, and with the transport of newly synthesized phospholipids, which could not be transported to the outer membrane outside of intact cells (149). Further studies utilized lipid A modifying enzymes to demonstrate that LPS accumulated at the periplasmic face of the inner membrane when the soluble periplasmic protein LptA was inactivated thus identifying a putative LPS carrier for this process (137). The idea of a protein bridge was further supported by crystal structures of LptA (145). The crystal structures revealed that LptA is primarily comprised of two 8-stranded β-sheets that fold into a β-jellyroll with a hydrophobic cleft at one edge (Figure 1.15). Moreover, in the presence of LPS (though LPS is not visible in the electron density), LptA crystallized as a chain of head-to-tail interacting subunits that formed two continuous β-sheets and a continuous hydrophobic groove suggestive of a hydrophobic slide for LPS (145).

Next, the crystal structure of the LptC periplasmic domain showed a very similar structure to that of LptA (Figure 1.15) supporting the model of a continuous β-jellyroll slide for LPS (146). Moreover, LptA could abstract LPS from LptC but not vice versa thus supporting affinity-based directional transport between proteins in the Lpt system though these results are also consistent with a soluble carrier mechanism of LptA (146). However, the stable protein bridge model was supported by the ability of the cytosolic protein LptB and inner membrane proteins LptC and LptF to pull down LptA and the outer membrane proteins LptD and LptE and by the ability of LptA to cause the association of liposomes containing the LptBCFG complex with liposomes containing the LptDE complex (138, 140).

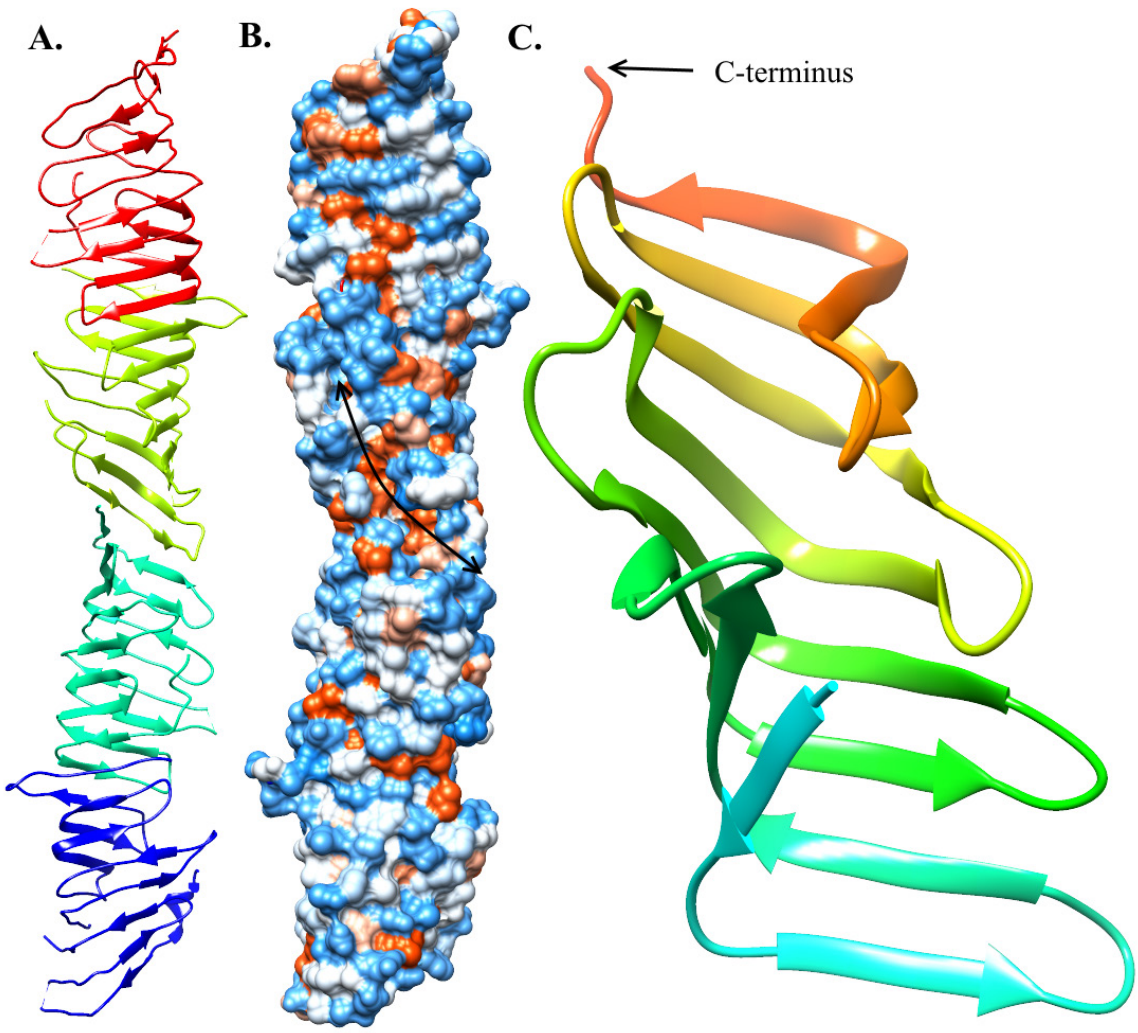


Figure 1.15: LptA and LptC. A. LptA tetramer with subunits interacting in a head-to-tail fashion (PDB: 2R1A) (145). The tetramer is shown with the N-terminal end at the bottom. **B.** Hydrophobicity surface of LptA tetramer in **A.** Hydrophobicity is show on an orange-blue scale with blue as the hydrophilic end. The arrow indicates the continuous, helical, hydrophobic groove (145). **C.** Ribbon structure of LptC periplasmic domain (PDB: 3MY2) (146).

Interactions between the Lpt proteins were further characterized by site-specific crosslinking with the non-natural amino acid p-benzoyl-L-phenylalanine (Bph) (143). Crosslinking suggested that the C-terminal end of the LptC β-jellyroll interacted with the N-terminal end of the LptA β-jellyroll, and the C-terminal end of the LptA β-jellyroll interacted with the N-terminal end of the LptD β-jellyroll (143). In addition, C-terminal

mutants of the LptC β-jellyroll were found to be nonfunctional *in vivo* and unable to pull down LptA, LptD, or LptE (141, 142). On the other hand, the N-terminal transmembrane domain of LptC (a predicted transmembrane helix) seemed to be of limited functional importance as it could be functionally replaced by unrelated transmembrane helices (141).

Crosslinking with Bph was also utilized to provide evidence for binding and directional transfer of LPS (139, 140). LPS was cross-linked to LptA and LptC when Bph was incorporated in the hydrophobic cleft of their β-jellyrolls but not when Bph was incorporated on the outer surface of their β-jellyrolls thus supporting transport along a continuous hydrophobic slide (139, 140). In addition, ATP-dependent transfer of LPS from LptBFG to LptC and LptA was observed *in vivo* and in liposomes, and transfer to LptA was enhanced by the inclusion of LptC (139, 140). Furthermore, LPS was transferred from liposomes containing LptBCFG to liposomes containing LptDE in an ATP- and LptA-dependent manner (140). Moreover, increase in ATP concentration resulted in LPS crosslinking to LptD at earlier time points and decreased crosslinking to LptD at later time points, which strongly indicated that transfer through the Lpt system was an active process that could be accelerated by an increased rate of ATP hydrolysis by LptB (140). Therefore, these data support a model wherein each new LPS molecule pushes the previous LPS molecules along the hydrophobic slide of the Lpt complex until they reach the outer leaflet of the outer membrane (140).

Crystal structures of the $LptB_2FG$ ABC transporter and of the LptB ATPase dimer bound to ATP have been solved and provides some insight into how ATP hydrolysis drives LPS through the Lpt system (144, 148, 150).

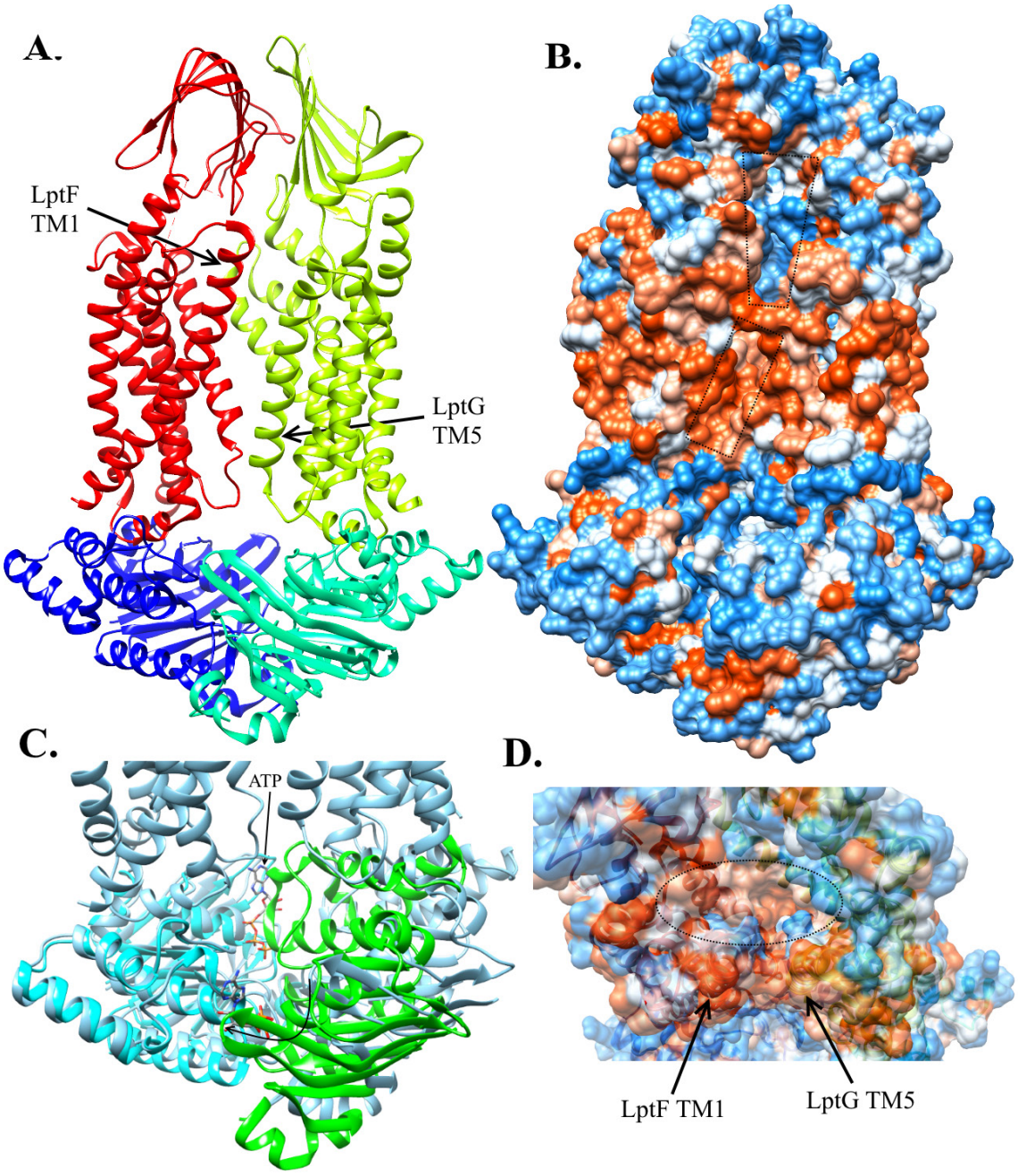


Figure 1.16: LptB$_2$FG ABC transporter. A. Ribbon structure LptB$_2$FG ABC transporter (PDB: 5X5Y) (144). LptF is shown in red while LptG is shown in yellow, and LptB subunits are shown in cyan and blue. **B.** Hydrophobicity surface of **A.** Hydrophobicity is shown on an orange-blue scale with orange as the most hydrophobic. Quadrilaterals mark spaces between LptF TM1 and LptG TM5. **C.** Comparison of ATP-bound LptB dimer (PDB: 4P33) (150) with LptB subunits in nucleotide-free LptB$_2$FG complex from **A.** Subunits from ATP-bound LptB-E163Q dimer are colored cyan and green, and the LptB$_2$FG complex is colored light blue. Cyan LptB-E163Q was overlaid with LptB chain B of complex (25). Curved arrow

shows movement of other LptB subunit from nucleotide-free conformation in $LptB_2FG$ complex to ATP-bound conformation in LptB-E163Q dimer. **D.** Rotated view of **B.** showing opening to V-shaped pocket on top/periplasmic side of $LptB_2FG$ complex. Ellipse marks the opening.

The structure of the $LptB_2FG$ complex was solved in the nucleotide-free state and shows that the transmembrane helices of LptF and LptG form a V-shaped pocket at the periplasmic face of the inner membrane (Figure 1.16) (144, 148). There are minimal contacts at the interface between TM1 of LptF and TM5 of LptG suggesting that these helices may separate to allow LPS to enter by lateral diffusion (Figure 1.16) (144, 148). The V-shaped pocket is lined by highly conserved hydrophobic residues that were found to be important for function *in vivo* (144). In addition, LptF and LptG have periplasmic β-jellyroll domains similar to those of LptA, LptC, and LptD (144–148). Mutation of the C-terminal strands of the LptF or LptG β-jellyroll domains caused growth defects in *E. coli*, which suggested that these domains may interact with the N-terminal end of the LptC β-jellyroll to form the first unit in the hydrophobic slide (148).

The LptB-E163Q dimer was crystallized with ATP bound showing that ATP binds between the ATPase subunits (Figure 1.16) (150). E163 is the predicted catalytic base, and mutation to Gln eliminated ATPase activity (150). Comparison of the ATP-bound LptB dimer and the nucleotide-free LptB dimer of the $LptB_2FG$ ABC transporter suggests that ATP binding to the transporter will cause the LptB subunits to rotate ~15^o (144, 150). Interactions of LptB with TM1 and the short α-helix between TM2 and TM3 of LptF and LptG likely serve to transmit this motion into the further opening of the space between LptF TM1 and LptG TM5 to allow LPS binding (144). Other conformational changes associated with the catalytic cycle presumably act to push LPS into the β-

jellyroll domain of LptF or LptG, but more work is required to elucidate this mechanism (144).

Finally, crystal structures of the LptDE outer membrane complex increase our understanding of how LPS is transported across the outer membrane and inserted into the extracellular leaflet (147, 151). The structures reveal that LptD forms a 26-stranded β-barrel, and LptE forms a 4-stranded β-sheet and 2 α-helices (Figure 1.17) (147, 151, 152). LptE acts as a plug for this β-barrel, and the extracellular side of the β-barrel is occluded by LptE and the extracellular loops of LptD (147, 151). On the other hand, the periplasmic side of the β-barrel is open to reveal a large, hydrophilic lumen that may accommodate the polysaccharide of LPS (Figure 1.17) (147, 151). LptE has been shown to bind LPS, and conserved basic residues of LptE that were found to be important for LPS binding are located at the extracellular side of the LptDE complex (144, 151, 152). Therefore, movement of LptE upon LPS binding may allow these residues to interact with the polysaccharide, and this interaction with LptE may help feed the polysaccharide through the β-barrel (147, 151, 152). The LptDE structures also revealed that the interactions of the first and final strands in the β-barrel are a weak point in the barrel structure thus suggesting that these strands may separate to allow the lipid A moiety of LPS to enter the membrane (147, 151). As mentioned above, LptD also has an N-terminal, periplasmic β-jellyroll domain that likely forms the final subunit in the hydrophobic slide, and this domain was crystallized with a detergent molecule bound in its hydrophobic cleft (Figure 1.17) (143, 147). This domain is located near the weak point in the β-barrel (147).

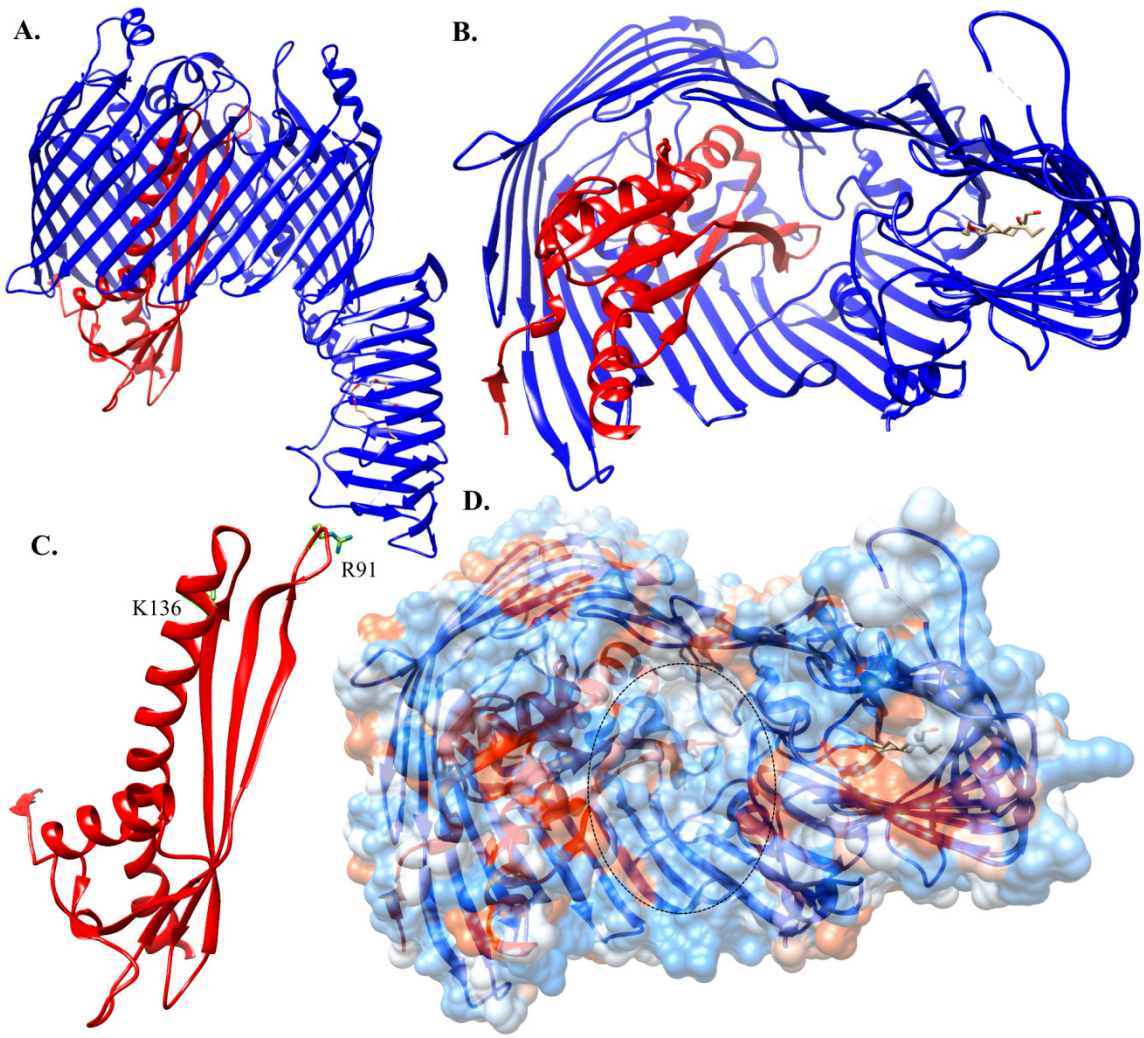


Figure 1.17: LptDE. A. Ribbon structure of LptDE complex (PDB: 4Q35) (147). LptD is colored blue, and LptE is colored red. A C8E detergent molecule is shown bound to the N-terminal β-jellyroll domain of LptD. **B.** Periplasmic view of **A. C.** LptE from the LptDE complex in **A.** Conserved basic residues shown to be important for LptE binding to LPS are highlighted (152). **D.** Hydrophobicity surface of **B.** Hydrophobicity is shown in an orange-blue scale with orange as the hydrophobic end. The ellipse marks the periplasmic entrance to the LptDE lumen.

1.5 Conclusions and Research Goals

This chapter has described the role of LPS in bacterial physiology and human disease, detailed how the structure of this important glycolipid was determined, and summarized

what is known about how LPS is synthesized and transported to the outer leaflet of the outer membrane. Throughout this chapter, the structures of the enzymes in the Raetz pathway, which produce the most conserved portion of LPS, and other important proteins involved in the synthesis and transport of LPS were highlighted (1). The discussion of these structures has served to demonstrate the great insight into biological processes that protein structures can provide, particularly when combined with additional biochemical analyses. Thus, the goals of my thesis research have been to solve and further characterize the structures of two of the enzymes in the Raetz pathway, LpxH and LpxB. The primary goal of solving the structures was accomplished by x-ray crystallography. The enzymology of LpxB and the enzymology and structural dynamics of LpxH were further explored. My research on each of these enzymes is detailed in the following chapters.

CHAPTER 2: STRUCTURAL DYNAMICS OF THE SUBSTRATE-BINDING CAPPING HELICES OF THE UDP-DIACYLGLUCOSAMINE PYROPHOSPHATASE, LpxH

Content in this chapter is reproduced, as per journal policy http://www.jbc.org/site/misc/edpolicy.xhtml#pip, from the article published in the Journal of Biological Chemistry: Bohl, T. E., Ieong, P., Lee, J. K., Lee, T., Kankanala, J., Shi, K., Demir, Ö., Kurahashi, K., Amaro, R. E., Wang, Z., and Aihara, H. (2018) The substrate-binding cap of the UDP-diacylglucosamine pyrophosphatase LpxH is highly flexible, enabling facile substrate binding and product release. J. Biol. Chem. 293, 7969–7981

2.1 Background

Early experiments indicated that hydrolysis of the phospho-anhydride of UDP-2,3-bis[O-(3R)-3-hydroxymyristoyl]-glucosamine (UDP-DAG) is the fourth step in the Raetz pathway (57, 153). The responsible pyrophosphatase in *E. coli* (LpxH) was identified as an essential Mn^{2+}-dependent hydrolase that catalyzes the attack of water on the α-phosphate of UDP-DAG (81, 84, 86). Electron paramagnetic resonance was utilized to further characterize Mn^{2+} binding, and these data were consistent with the presence of a binuclear manganese center (84).

Recently, LpxH was added to the list of structurally characterized enzymes in the Raetz pathway with the publication of structures of LpxH from *Haemophilus influenzae* and *Pseudomonas aeruginosa* bound to the 2,3-bis[(3R)-3-hydroxymyristoyl]-glucosaminyl-1-phosphate (lipid X) product and in the unbound form (85, 154). LpxH has a Calcineurin-like metal-dependent phosphoesterase fold with the addition of a unique helical cap, comprising 3 major α-helices, that covers the active site and binds the lipid substrates (Figure 2.0) (85, 154). The core domain is composed of 2 five-stranded β-sheets surrounded by 4 α-helices, and the helical cap domain is inserted between β-strand 6 and the final α-helix (85, 154). The crystal structures confirm that LpxH has a binuclear manganese center in its active site (85, 154). In the *P. aeruginosa* structure (PDB: 5B49), the first Mn^{2+} is coordinated by D8, H10, D41 (bridges the Mn^{2+}), H197, a bridging water, and a second water that completes the octahedral geometry and may represent the attacking water (85). Alternatively, the bridging water may act as the attacking water (155). The second Mn^{2+} is coordinated by D41, the bridging water, N79, H114, and H195, which leaves one open coordination site that may be filled by the α-phosphate in

the substrate-bound structure (Figure 2.0) (85, 155). The importance of several of the Mn^{2+}-coordination residues was confirmed by enzymatic assays in *H. influenzae* LpxH (HiLpxH) that showed large (3-5 orders of magnitude) drops in specific activity when individual residues were mutated to Ala (84).

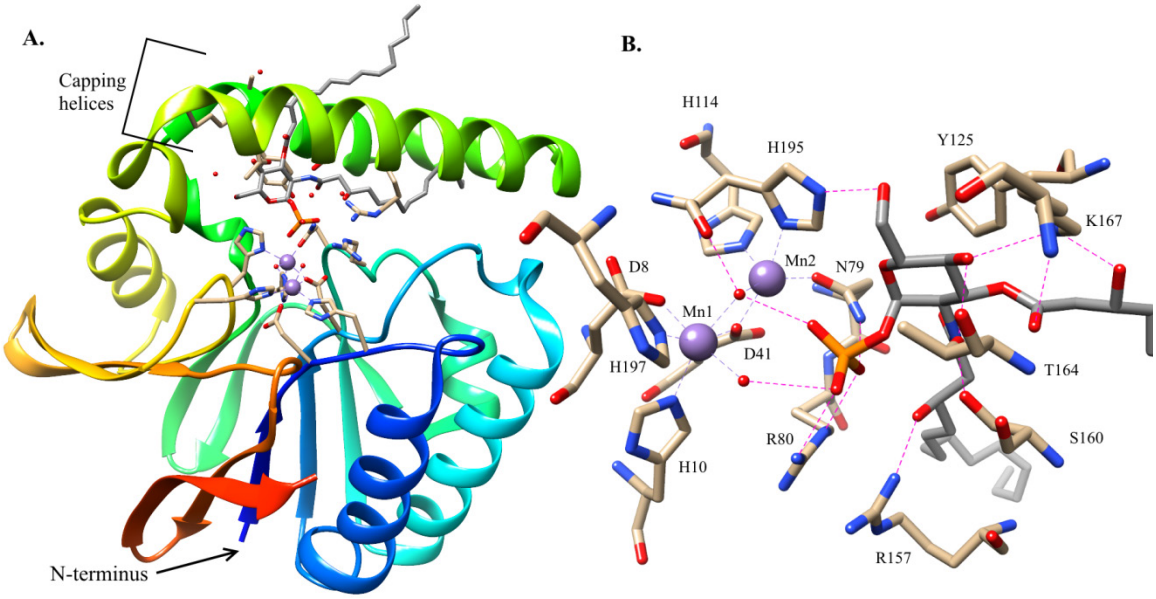


Figure 2.0: *Pseudomonas* LpxH bound to lipid X product. A. Ribbon structure of PaLpxH bound to lipid X (PDB: 5B49) (85). Lipid X carbons are colored grey, and protein carbons are colored tan. Mn^{2+} are colored purple. Spectrum coloring begins with blue at the N-terminus. **B.** Residues involved in lipid X binding from **A.** Metal coordination is shown in purple with the following distances in Ångstroms: from Mn301 to H10 Nε (2.3), to H197 Nε (2.3), to D8 Oδ1 (2.2), to D41 Oδ2 (2.3), to bridging water (2.3), and to adjacent water (2.2); and from Mn302 to D41 Oδ2 (2.2), to N79 Oδ (2.1), to H195 Nδ (2.3), to H114 Nε (2.2), and to bridging water (2.2). Hydrogen-bond and salt-bridge interactions are shown in pink with distances in Ångstroms of 2.7 from H195 O to bridging water, 2.9 from 2-acyl chain β-hydrxyl to R157 Nη1, 2.8 from glucosamine 4-hydroxyl to T164 Oγ and 2.9 to K167 Nζ, 3.3 from K167 Nζ to 3-acyl chain carbonyl and 2.8 to its β-hydroxyl, 2.8 from glucosamine 6-hydroxyl to H195 Nε, 2.8 from glucosamine 2-amino to S160 Oγ, and 2.7 from 1-phosphate to bridging water and 3.4 to adjacent water and 2.6 to N79 Nδ and 2.8 and 3.1 respectively to R80 Nε and Nη2.

Lipid X-bound structures show that the lipid X head group is recognized by extensive hydrogen-bonding interactions, and the acyl chains bind in the hydrophobic pocket formed by the capping domain and the interface with the core domain: the amide-

linked chain is buried in the hydrophobic pocket while the ester-linked chain exits between the helices and extends across the top of the lid or into solvent (85, 154). In the *P. aeruginosa* LpxH (PaLpxH) structure (PDB: 5B49), the lipid X phosphate is bound by N79 and R80, which is a conserved residue particular to LpxH and is usually His in other metal-dependent phosphoesterases of the Calcineurin-like superfamily (85, 155). Consistent with the importance of this Arg, the Ala mutation in HiLpxH decreased activity 7000-fold (84). In addition, glucosamine is hydrogen-bonded to S160, T164, K167, and H195, and the β-hydroxyl groups of the acyl chains are hydrogen-bonded to R157 and K167 (Figure 2.0) (85). Finally, Y125 packs on top of the glucosamine ring (85). All of these residues noted from PaLpxH are conserved in *E. coli* LpxH (same numbering) and HiLpxH (numbering +1) with the exception of T164, which is Asn and Lys, respectively.

While previous structures suggested that the capping domain is stably attached to the rest of the protein and only becomes slightly disordered in the absence of ligand (85), we provide multiple lines of evidence that the capping domain is highly dynamic in the absence of ligand. These results have implications for the mechanism of substrate binding and product release: although slight disordering of the third capping helix as observed previously could allow lipids to enter and exit the active site, a more extensive conformational change would enable facile binding and release of lipids through a wider opening between the capping helices.

2.2 Results and Discussion

Crystal Structure

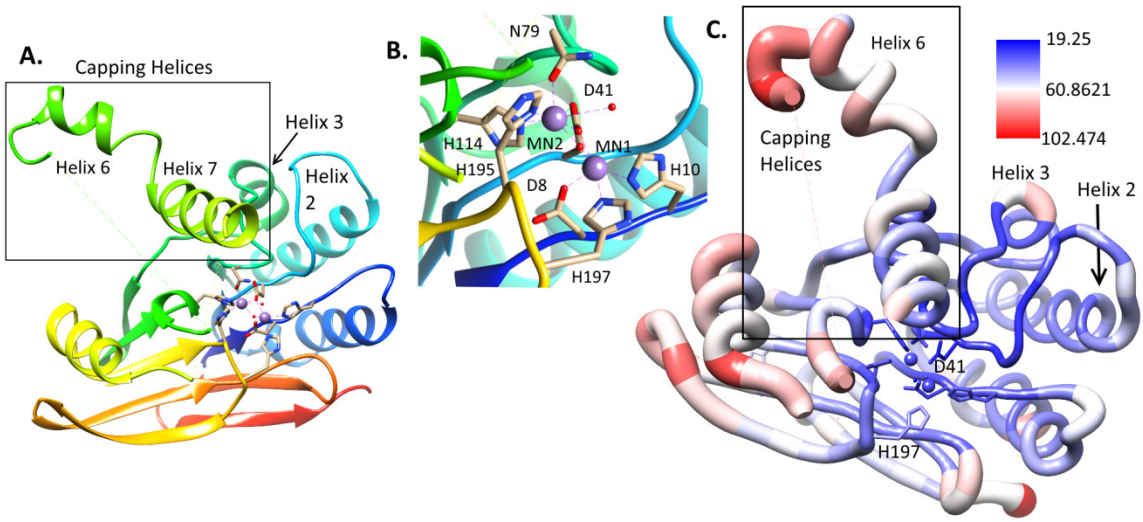


Figure 2.1: *E. coli* LpxH crystal structure. A. The overall structure of LpxH4+4 shows a Calcineurin-like metal-dependent phosphoesterase fold with two Mn^{2+} bound in the active site and the partially resolved capping helices displaced from the core domain. Spectrum coloring begins with blue at the N-terminus. **B.** Two Mn^{2+} are coordinated in the active site with the following distances in Ångstroms: D8 Oδ1 to Mn1 (2.1), H10 Nε to Mn1 (2.2), D41 Oδ2 to Mn1 (2.3) and to Mn2 (2.1), N79 Oδ to Mn2 (2.1), H114 Nε to Mn2 (2.1), H195 Nδ to Mn2 (2.2), H197 Nε to Mn1 (2.2), bridging water to Mn1 (1.9) and to Mn2 (2.1), and adjacent water to Mn2 (3.3). **C.** LpxH4+4 rendered by average B-factor. Wider tubes and colors closer to red indicate higher B-factors. Except where otherwise noted, comparison of structures and preparation of structures for figures was performed in UCSF Chimera (25).

A solubilized version of *E.* coli LpxH including 4 solubilizing mutations (F141H, L142S, L146S, and F147H) and 4 surface entropy reduction mutations (E14A, E15A, K161T, and E162A) (LpxH4+4) was generated to improve protein expression and crystallization. LpxH4+4 crystallized in plate-shaped crystals that diffracted to 2.00 Å, and the structure was solved by molecular replacement (Table 2.1) (PDB: 5WLY). The core phosphoesterase domain of *E. coli* LpxH is composed of two 5-stranded β-sheets sandwiched together by α-helices, and this domain overlays very well with the previous LpxH structures (25, 85, 154) (Figures 2.1A and 2.2A).

Table 2.1: Diffraction and Refinement Statistics

Parameter	LpxH4+4 (PDB: 5WLY)	LpxH4+4 SeMet (SAD data)
Diffraction		
Wavelength (Å)	0.97918	0.97918
Resolution range (Å)	66.05-2.00 (2.06-2.00)	66.08-2.63 (2.75-2.63)
Space group	P 21 2 21	P 21 2 21
Unit cell a,b,c (Å) α,β,γ (°)	57.012 62.064 66.047 90 90 90	56.66 62.77 66.08 90 90 90
Total observations	96359 (6401)	49232 (6109)
Total unique	16046 (1090)	7415 (890)
Multiplicity	6 (5.9)	3.6 (3.9)
Completeness (%)	98.6 (91)	99.5 (99.7)
Mean I/σI	7.7 (0.8)	10.3 (1.2)
Wilson B-factor	26.88	49.49
R-merge	0.179 (1.973)	0.198 (1.932)
R-meas	0.196 (2.162)	0.232 (2.257)
R-pim	0.078 (0.867)	0.120 (1.160)
$CC_{1/2}$	0.994 (0.279)	0.992 (0.395)
Refinement		Not fully refined
Resolution Range (Å)	45.23-2.00 (2.16-2.00)	
Completeness (%)	90.44 (71.09)	
Reflections	14750 (2414)	
R-free Reflections	689 (117)	
R-work/R-free	0.1983/0.2271 (0.3310/0.3626)	
Non-hydrogen atoms	1850	
Protein	1687	
Ligands/ions	10	
Solvent	153	
RMS bond lengths (Å)	0.003	
RMS bond angles (°)	0.594	
Ramachandran favored (%)	94.86	
Ramachandran allowed (%)	4.67	
Ramachandran outliers (%)	0.47	
Rotamer outliers (%)	1.80	
Cβ-outliers	0	
Clashscore	2.38	
Average B-factor	42.45	
Protein	42.46	
Ligands/ions	48.43	
Solvent	41.97	
Number of TLS groups	3	

Statistics for the highest resolution shell are shown in parentheses. Each dataset was collected with a single crystal.

However, the capping helices of *E. coli* LpxH were displaced from the position observed in previous structures and were largely disordered (Figure 2.2A). Residues 122-130 and 162-172 from the capping helices are not visible. While the unbound form of *P. aeruginosa* LpxH (PaLpxH) did show increased disorder and an altered conformation where the capping helices bind the product head group and connect to the core domain, the rest of the helices remained packed on top of the active site (Figure 2.2D) (85). In *E. coli* LpxH, the helices are detached from the active site, and only the middle portion (residues 131-161) is visible. This conformation is fortuitously stabilized by crystal contacts (Supplementary Fig. 2.1) and is likely transient in solution. However, this structure suggests that the capping helices are much more flexible and disordered in the absence of substrate than was apparent from previous LpxH structures. This structure of *E. coli* LpxH was the impetus for further experiments to examine the movement of the cap. It is important to note that the capping helices of LpxH4+4, which contains 6 mutations within the capping helices, or of *E. coli* LpxH in general may be inherently more flexible than those in *H. influenzae* or *P. aeruginosa* LpxH. On the other hand, in previous LpxH crystal structures, the closed conformation was stabilized by binding to lipid X and/or crystal contacts wherein the capping domains were often packed against each other, and this may explain the limited movement observed in the apo PaLpxH structure (85, 154). Because the structures and sequences of LpxH from these species are highly similar (Figures 2.2 and Supplementary Fig. 2.2), the design and analysis of the following experiments were guided by the hypotheses that the motions of the capping helices are similar and that the conserved residues play the same roles in these different LpxH variants.

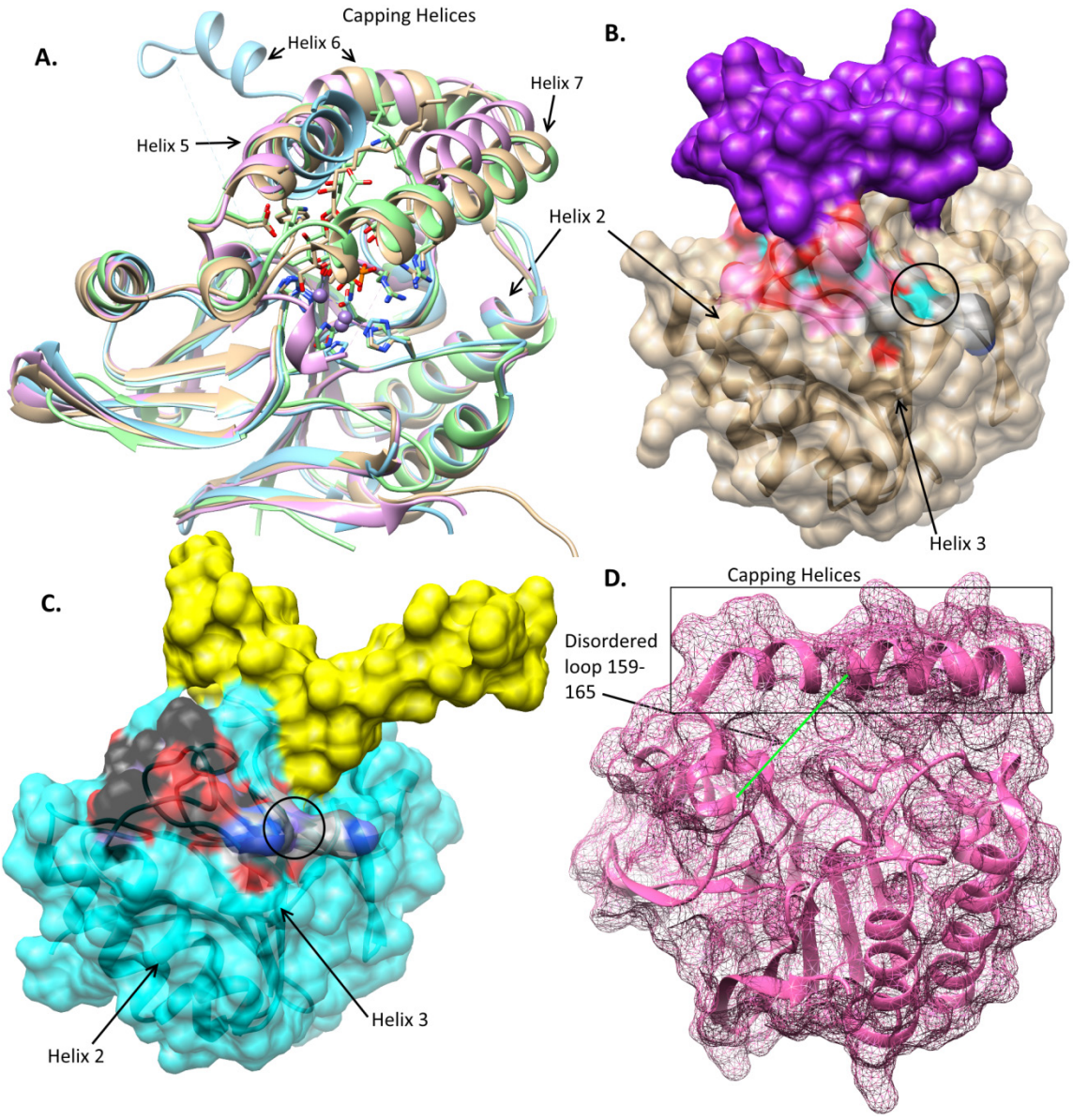


Figure 2.2: Comparison with HiLpxH and PaLpxH. A. Overlay of LpxH4+4 (PDB: 5WLY, blue) with HiLpxH (PDB: 5K8K, green), PaLpxH with no Mn^{2+} (PDB: 5B4A chain B, tan), and apo PaLpxH (PDB: 5B4D chain A, pink). The RMSD values between the core domains are as follows: 0.784 Å between 180 atom pairs for 5WLY and 5B4A, 0.782 Å between 181 atom pairs for 5WLY and 5B4D, and 0.730 Å between 164 atom pairs for 5WLY and 5K8K. While the core domains overlay well, the capping helices in LpxH4+4 are displaced from their position in the closed structures, and helix 5 is not visible in the electron density. In the apo structure of PaLpxH, the third helix of the capping domain is partially disordered and slightly above its position in the lipid X-bound structures, but the displacement is much less than in the LpxH4+4 structure. **B.** The surface rendering of PaLpxH bound to Mn^{2+} and lipid X (PDB: 5B49) shows that the capping helices (purple) are closed down on the core domain, but residues 84-88 (grey carbons),

which, in LpxH4+4, underwent hydrogen-deuterium exchange more rapidly than any part of the protein outside of the capping helices (Figures 2.3 and 2.4), remain largely exposed. In contrast, the loop between β-strand 2 and α-helix 2 (pink carbons) is largely covered by the cap in this conformation. The backbone nitrogens of these two loops are colored cyan, and the exposed nitrogens of residues 86 and 87 are circled. **C.** The surface rendering of LpxH4+4 shows that movement of the capping helices (yellow) exposes a large area of the core domain including the loop between strand 2 and helix 2 (black carbons). However, the exposure of residues 84-88 (grey carbons) is similar to that in the lipid X-bound conformation of PaLpxH. The backbone nitrogens of these loops are colored purple, and the exposed nitrogens of residues 86 and 87 are circled. **D.** The surface rendering of apo PaLpxH (PDB: 5B4D) shows that the capping domain remains packed against the core domain even when with the partial disordering of helix 7.

Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS)

HDX-MS of LpxH4+4 showed that the capping domain has a relatively high exchange rate compared to the rest of the protein (Figures 2.3 and 2.4). Residues 146-156, which constitute the first half of the third helix of the capping domain (helix 7), have the highest exchange rate (66.7% in 1 min) of any detectable portion of the protein (Figures 2.3 and 2.4). Residues 158-172 have the second highest exchange rate (53.8% for residues 158-169 and 55.9% for residues 170-172 in 1 min) and form the second half of helix 7 and the C-terminal loop of the capping domain. In the LpxH4+4 structure, only residues 158-161 were visible from this second region, and residues 159-165 were disordered in the apo PaLpxH structure (Figures 2.1 and 2.2) (85). Residues 84-88, which form a loop and the N-terminus of helix 3, have the highest exchange rate outside of the capping domain (47.0% in 1 min) (Figures 2.3 and 2.4). As these residues are largely surface exposed even in closed structures (85, 154), this provides little insight (Figure 2.2). Unfortunately, the residues in the long loop between sheet 2 and helix 2, which are covered by the free end of the capping domain in the closed structures but surface-exposed in our structure, were not observed by mass spectrometry (Figures 2.2 and 2.3).

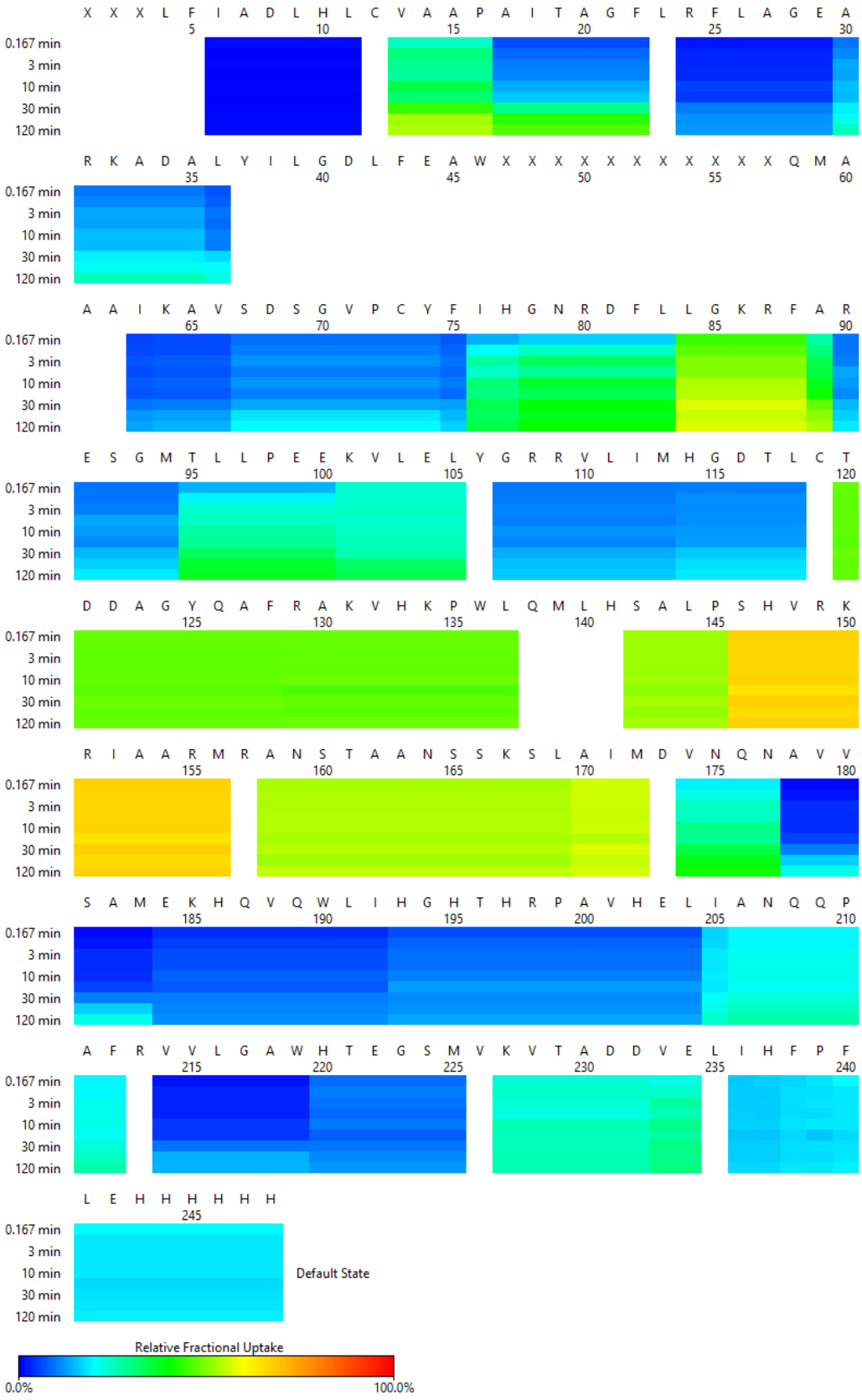

X X X L F I A D L H L C V A A P A I T A G F L R F L A G E A
5 10 15 20 25 30
R K A D A L Y I L G D L F E A W X X X X X X X X X X Q M A
35 40 45 50 55 60
A A I K A V S D S G V P C Y F I H G N R D F L L G K R F A R
65 70 75 80 85 90
E S G M T L L P E E K V L E L Y G R R V L I M H G D T L C T
95 100 105 110 115 120
D D A G Y Q A F R A K V H K P W L Q M L H S A L P S H V R K
125 130 135 140 145 150
R I A A R M R A N S T A A N S S K S L A I M D V N Q N A V V
155 160 165 170 175 180
S A M E K H Q V Q W L I H G H T H R P A V H E L I A N Q Q P
185 190 195 200 205 210
A F R V V L G A W H T E G S M V K V T A D D V E L I H F P F
215 220 225 230 235 240
L E H H H H H H H
245
0.167 min
3 min
10 min
30 min
120 min
Default State
Relative Fractional Uptake
0.0%
100.0%

Figure 2.3: Time-dependent hydrogen-deuterium exchange. Heat map of fraction of deuterium uptake plotted against the amino acid sequence at 0.167, 3, 10, 30, and 120 min. Gaps show where no uptake data were available because the amino acid was not observed or because the amino acid was only observed at the N-terminus of peptides. Uptake is most rapid for residues 146-156 followed by residues 158-172; both of these regions are part of the helical cap domain. The most rapid uptake outside of the cap is observed for residues 84-88, which are largely surface exposed in our structure and in the closed structures (85, 154).

Nonetheless, the high exchange rate of the capping domain, particularly in the 146-156 region that was ordered and attached to the catalytic core in previous structures, supports the hypothesis that this domain is highly dynamic rather than stably attached to the core domain.

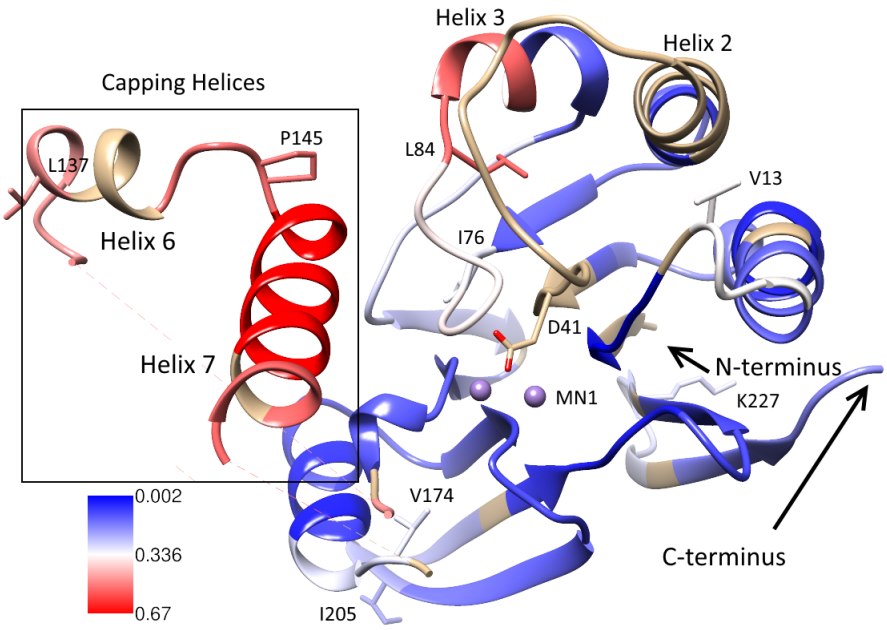


Figure 2.4: Fractional exchange at 10 min. The 10 minute fractional deuterium uptake values plotted onto our *E. coli* LpxH structure (coloring with low blue end and high red end). Regions with no exchange data are colored tan. Active site Mn^{2+} ions are shown in purple.

Molecular Dynamics Simulations Show Cap Flexibility

PaLpxH was selected for molecular dynamics (MD) simulations because structures of PaLpxH have been determined with and without lipid X bound (85). Moreover, simulations with *E. coli* LpxH would require a large portion of the cap to be

computationally built before molecular dynamics could be run; therefore, utilizing PaLpxH structures to minimize the amount of structure that needed to be computationally built gave more reliable starting models for MD simulations.

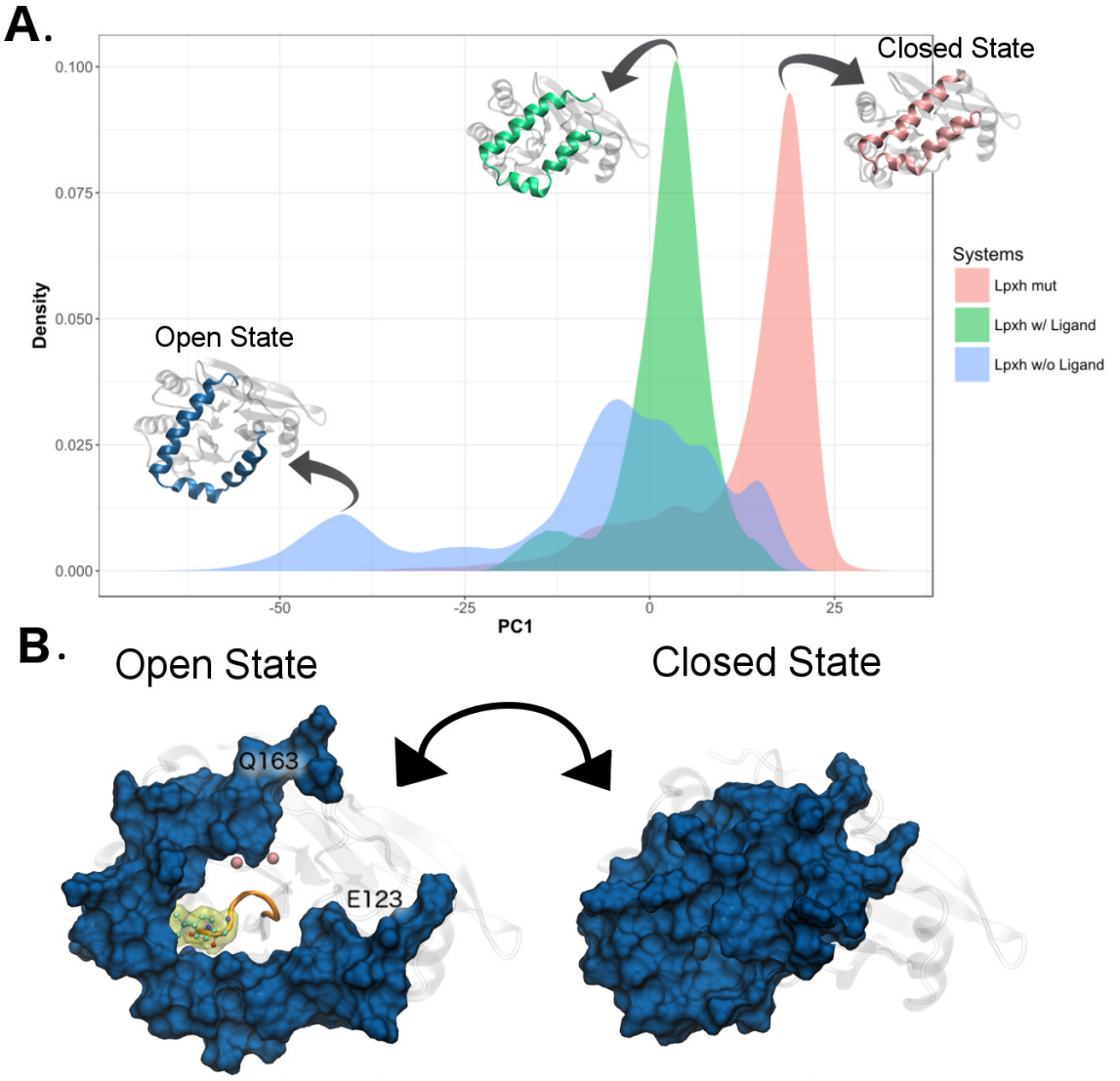


Figure 2.5: Flexibility of the capping helices. A. Density map of the first principal component is shown. **B.** The open and closed forms of the cap found in MD simulations are shown. Cap residues (121 to 168) are colored in blue while the rest of the protein is shown in silver surface. Mn^{2+} ions are shown in pink. The loop containing F82 and L83 is colored orange, and the residues are colored with green carbons and a yellow surface.

MD simulations were performed on PaLpxH systems with lipid X bound (PDB: 5B49) ($LpxH_{holo}$), with lipid X removed (PDB: 5B49) ($LpxH_{apo\_holo}$), and with the unbound structure (PDB: 5B4C) ($LpxH_{apo\_apo}$) (85). In our 3 μs simulations of $LpxH_{holo}$, lipid X conserved its conformation in the active site. Principal component analysis showed that the cap mostly sampled the conformation observed in the crystal structure and rarely sampled other conformations with the cap slightly more open. In the presence of lipid X, the cap adopted the crystal structure conformation with only small deviations (Figure 2.5A). In the absence of a ligand in PaLpxH ($LpxH_{apo\_apo}$ and $LpxH_{apo\_holo}$), the cap became more flexible, sampling many conformations including an open configuration where residues 123 and 163 were farther apart (Figure 2.5A). $LpxH_{apo\_apo}$ displayed an open cap where a large cavity formed near the Mn^{2+} active site that is large enough to accommodate the phosphate and glucosamine moieties of lipid X (Figure 2.5B). Overall, MD simulations agreed well with experimental results in that the cap is rather flexible in the absence of lipid X.

The Role of Loop (Residues 80 to 83) in Cap Flexibility

In our MD simulations of apo LpxH starting from the lipid X-bound conformation ($LpxH_{apo\_holo}$), we observed that the cap generally maintained its positioning as observed in the PaLpxH crystal structures. However, as a notable deviation from the crystal structure we found that residues 80 to 83, which formed a flexible loop underneath the cap, move upward to and protrude into the cap. In particular, F82 and L83 side chains interact with the cap residues via hydrophobic interactions. To further investigate the role of the loop, we simulated a F82G/L83G double mutant of PaLpxH ($LpxH_{mut}$). Principle component analysis of the $LpxH_{mut}$ showed that the mutations caused the cap to collapse

on itself, adopting a closed conformation (Figure 2.5). Based on this, the loop may have a role in regulating the opening or closing of the cap and thus the exposure of the active site for ligand binding.

We then studied the root-mean-square-deviation (RMSD) of the loop for individual systems to elucidate the dynamics of the loop. We aligned the trajectories against two reference conformations (loop-down and loop-up conformations) to identify which conformation the loop adopts over time. The loop-down conformation has the loop farther away from the cap as in the crystal structure, while the loop-up conformation has the loop protruded inside the cap as seen in simulations (Figure 2.6A). In both $LpxH_{mut}$ and $LpxH_{holo}$ systems, the loop adopted the down conformation away from the cap, while the loop in the $LpxH_{apo\_apo}$ and $LpxH_{apo\_holo}$ systems sampled both the up and down conformations and conformations in-between (Figure 2.6A).

To test whether the loop-up and loop-down conformations of the loop correlate with the open and closed states of the cap, we calculated the distance between the cap and the loop and the distance between residues 123 and 163, the two residues at the opposite ends of the cap (Figure 2.5B). These two residues provide a good measure of the cap's open and closed conformations. As seen in Figure 2.6B, the data indicated a negative correlation between the distance between residues 123 and 163 and the distance between the loop and the cap; thus, the loop-up conformation correlates with the open conformation. Taken together, our results suggest a mechanism where the loop protrudes into the cap to keep it open, forming a solvent-exposed cavity for ligand binding. Therefore, mutation of the residues 82 and 83 (F82G, L83G) would be expected to decrease activity by inhibiting substrate binding.

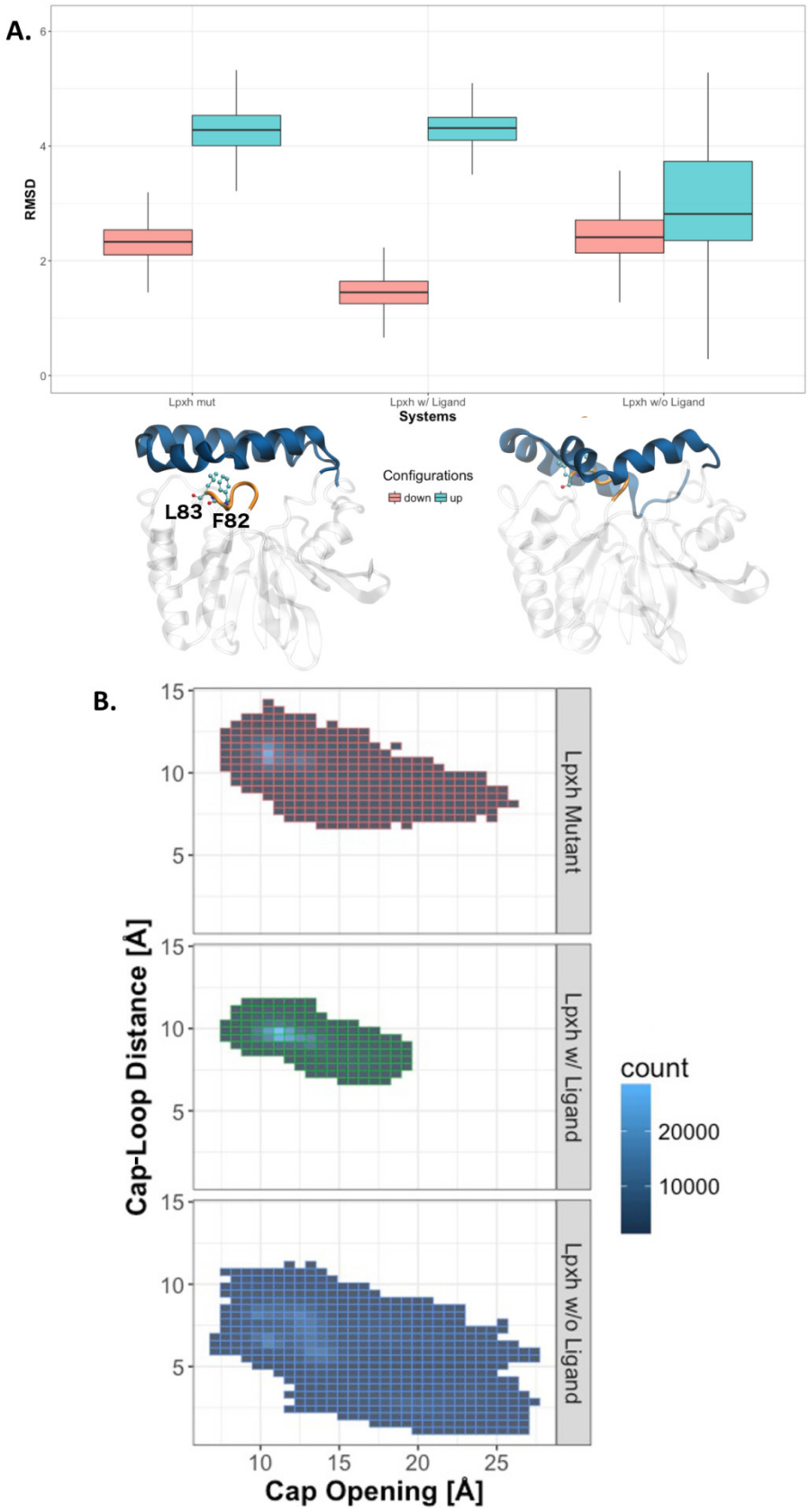

A.
RMSD
0
2
4
6
Lpxh mut
Lpxh w/ Ligand
Lpxh w/o Ligand
Systems
Configurations
down
up
L83
F82
B.
Cap-Loop Distance [Å]
Cap Opening [Å]
Lpxh Mutant
Lpxh w/ Ligand
Lpxh w/o Ligand
count
20000
10000

Figure 2.6: Correlation of loop movement and cap opening. A. A boxplot of the RMSD calculation aligned against the loop-up (teal) and loop-down (salmon) conformations. The loop-up and loop-down protein conformations in ribbon diagram are shown at the bottom on the right and left, respectively. The cap and the loop are highlighted in blue and orange, respectively, and F82 and L83 are shown with cyan carbons. From left to right, plots for $LpxH_{mut}$, $LpxH_{holo}$, and $LpxH_{apo\_holo}$/$LpxH_{apo\_apo}$ are shown.
B. Heatmap showing correlation of the cap opening with the cap-loop distance. The cap opening distance is measured between residues 123 and 163 (Figure 2.5) while the cap-loop distance is determined by the distance between the center of mass of the cap and the center of mass of the loop. Individual points represent frequency of conformations observed during the time-course of MD simulations. Lighter colors represent greater point density.

To test this hypothesis, these conserved residues were mutated to Gly in the soluble *E. coli* LpxH forming LpxH4+4-GG, and its activity was tested *in vitro* as described below.

Activity of E. coli *LpxH Mutants*

Activity assays showed that the solubilized *E. coli* LpxH crystallized in this study was catalytically competent for hydrolyzing UDP-DAG to lipid X (Figure 2.7 and Table 2.2). In fact, the crystallized form showed significantly more activity than the wild-type enzyme with a C-terminal mCherry fusion (LpxH-mCherry) (Table 2.2). These results indicate that none of the 8 mutated residues were important for activity on detergent-solubilized substrate. The reason for increased activity is unknown but may simply reflect improved solubility: less aggregated protein is present. Alternatively, the 6 mutations in the capping helices may increase the flexibility of the cap and thus facilitate cap opening. The LpxH3+4-W136H mutant, wherein F141 was restored and W136 was mutated as an alternate solubilizing mutation, has similar activity to LpxH4+4. In *H. influenzae* LpxH (HiLpxH), the corresponding F142 contacts the amide-linked acyl chain of lipid X while the residue corresponding to W136 is surface exposed in all LpxH structures (Figure 2.8) (85, 154). The similar solubility and activity of these variants is more consistent with the

hypothesis that increased activity reflects improved solubility of the LpxH variants rather than effects on substrate binding.

Table 2.2: LpxH Specific Activities

LpxH	Specific Activity ± Standard Error [$\mu mol*min^{-1}*\mu mol^{-1}$]	95% Confidence Interval	Percent of wtLpxH-mCherry (with DMSO)
-mCherry	12.75 ± 1.84	8.41 to 17.1	100%
4+4	62.17 ± 5.20	49.89 to 74.45	488%
3+4-W136H	76.54 ± 6.04	62.25 to 90.83	600%
4+4-CC	17.94 ± 2.81	11.28 to 24.59	141%
4+4-GG	1.438 ± 0.226	0.9039 to 1.971	11.3%
-mCherry (10% DMSO)	14.22 ± 0.68	12.62 to 15.81	112% (100%)
-mCherry (2 µM inhibitor)	6.019 ± 0.353	5.185 to 6.853	47.2% (42.3%)
-mCherry (6 µM inhibitor)	2.993 ± 0.140	2.661 to 3.325	23.5% (21.0%)
-mCherry (10 µM inhibitor)	1.803 ± 0.247	1.219 to 2.388	14.1% (12.7%)
-mCherry (50 µM inhibitor)	1.497 ± 0.0860	1.293 to 1.7	11.7% (10.5%)

Reactions were run in triplicate at ambient temperature (21^{o}C) with 100 µM UDP-DAG and 5 to 20 nM LpxH (in 0.5 M NaCl, 20 mM Tris-HCl pH 7.5, 5 mM DTT, and 2.5 mM $MnCl_2$), 0.1 M Tris-HCl pH 8.0, 0.1% Triton X-100, and 1 mg/mL BSA. Inhibitor (**3**) was dissolved in DMSO resulting in a final concentration of 10% DMSO in these reactions. Activity was measured with a UMP/CMP-Glo kit (Promega), which quantifes UMP concentration with a luciferase-coupled assay. Specific activities with standard errors and confidence intervals were caculated by linear regression analysis of 3 time points within the linear range of the reactions in Graphpad Prism v7.03.

However, mutating I47 and R149 to Cys (LpxH4+4-CC) decreased activity significantly placing it on par with LpxH-mCherry (Table 2.2): I47 forms part of the hydrophobic pocket that binds the hydrocarbon tails of the substrate, and R149 forms a salt-bridge with D50 in structures where the cap is positioned on top of the active site, not displaced

as in the LpxH4+4 structure (Figure 2.8). The loss of this salt-bridge that stabilizes the position of the cap may explain the decreased activity of LpxH4+4-CC. As described above, molecular dynamics simulations suggested that residues F82 and L83 are important for stabilizing an open cap conformation in which the active site is accessible to substrate. When both of these residues were mutated to Gly (LpxH4+4-GG), activity was decreased significantly below that of LpxH-mCherry consistent with the effect predicted by molecular dynamics simulations (Table 2.2). However, these mutations could also decrease substrate binding efficiency by decreasing the hydrophobic contacts available to the amide-linked acyl chain of lipid X. In addition, the melting temperature (T_m) of LpxH4+4-GG was decreased by 2.6^{o}C relative to LpxH4+4, indicating that these mutations partially compromise protein stability (Supplementary Table 2.1). However, the T_m remained well above the temperature (21^{o}C) at which activity assays were run. Finally, because even the most active LpxH variant tested here had a specific activity ~2500-fold lower than that measured for wild-type HiLpxH (84), it is important to note that our reaction conditions were not identical. In particular, we determined specific activities with reactions run at 21^{o}C versus 30^{o}C used previously (84).

Inhibition

Nayar *et al.* (156) identified an inhibitor of *E. coli* LpxH via genetic screening. Thus, we tested the ability of this inhibitor (**3**) (Scheme 2.1) to inhibit LpxH *in vitro*. Luminescence-based assays gave an IC_{50} of 1.2±0.2 μM for LpxH-mCherry (Table 2.2 and Supplementary Fig. 2.3). Thin layer chromatography (TLC) based assays also showed inhibition of LpxH-mCherry by **3**, giving incomplete conversion of UDP-DAG to lipid X (Figure 2.7): in TLC-based assays, the IC_{50} appeared to be low micromolar (~5

μM) (Figure 2.7). On the other hand, the crystallized form of *E.* coli LpxH was not inhibited by **3** (Figure 2.7).

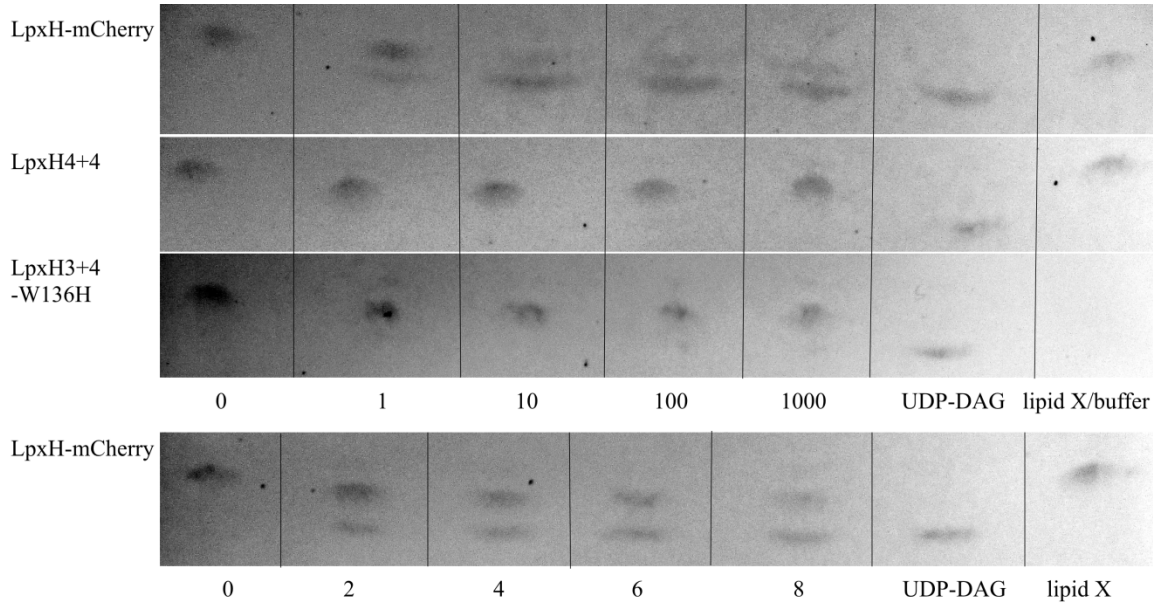


Figure 2.7: Inhibition of LpxH. Reactions (10 μL) were run at 30°C with 62 μM UDP-DAG, 10 nM LpxH, 0.1 M Tris-HCl pH 8.0, 0.1% Triton X-100, 10% DMSO, and 1 mg/mL BSA at various concentrations of the inhibitor (**3**) (Scheme 2.1) (shown in micromolar). Reactions were quenched by spotting on HPTLC Silica gel 60 plates, which were run with 25/15/4/2 chloroform/methanol/water/acetic acid mobile phase and charred with 20% sulfuric acid in ethanol and a heat gun. Reactions prepared with UDP-DAG but no enzyme act as a negative control and UDP-DAG standard. Independently purified lipid X product is also included as a standard. LpxH-mCherry reactions were run for 30 min. The reaction proceeded to completion in the absence of inhibitor, but is slightly inhibited by 1 μM inhibitor. At 2-8 μM inhibitor, the reaction proceeds to about 50% completion, and it is mostly inhibited at 10 μM and above. These plates were not repeated. LpxH4+4 reactions were run for 20 min. No inhibition was observed up to 1 mM inhibitor. This plate was replicated twice. LpxH3+4-W136H reactions were run for 40 min. A small amount of UDP-DAG remains at 100 μM and 1 mM inhibitor suggesting weak inhibition at high inhibitor concentrations. This plate was replicated thrice. Analysis of band intensities is presented in Supplementary Fig. 2.6, and the full plates are shown in Supplementary Fig. 2.7.

This result is likely due in part to the F141H mutation because *E. coli* LpxH-F141L was previously identified as a resistant mutant (156). However, when F141 was reintroduced in LpxH3+4-W136H, inhibition was still weak at inhibitor concentrations as high as 1

mM (Figure 2.7). Thus, L142, L146, and/or F147 are likely also involved in inhibitor binding, possibly by influencing the position of F141. Influence on the position of F141 may also explain the resistance of the LpxH-R149H mutant: the corresponding Arg packs against F142 in the HiLpxH structure (PDB: 5K8K) (Figure 2.8). The G48D and L84R resistance mutations likely function by introducing a charged residue into the hydrophobic pocket thus blocking inhibitor binding. The location of these mutations strongly suggest the inhibitor acts by competitively binding to LpxH and blocking binding of the hydrocarbon tails of the substrate (156) (Figure 2.8). When the inhibitor was docked into the HiLpxH active site with AutoDock Vina (157), it tended to occupy the same space as the amide-linked acyl chain (Figure 2.8D). In one compelling model, the inhibitor is engaged with the F142 (F141 in *E. coli* LpxH) by π-stacking, which could explain the resistance conferred by the otherwise conservative F141L mutation (Figure 2.8D) (156). The HiLpxH structure was taken to be the best candidate for molecular docking because PaLpxH already contains L141.

2.3 Conclusions and Future Directions

In this work, we have provided multiple lines of evidence that the substrate-binding cap of LpxH is more dynamic than was apparent from previously published crystal structures alone (85, 154). Molecular dynamics simulations showed how this dynamic nature of the cap can allow for facile substrate binding and product release between the helices of the cap. Furthermore, these simulations identified two residues that may act as a wedge to promote cap opening to allow substrate entry, and mutation of these residues was found to decrease LpxH activity as predicted.

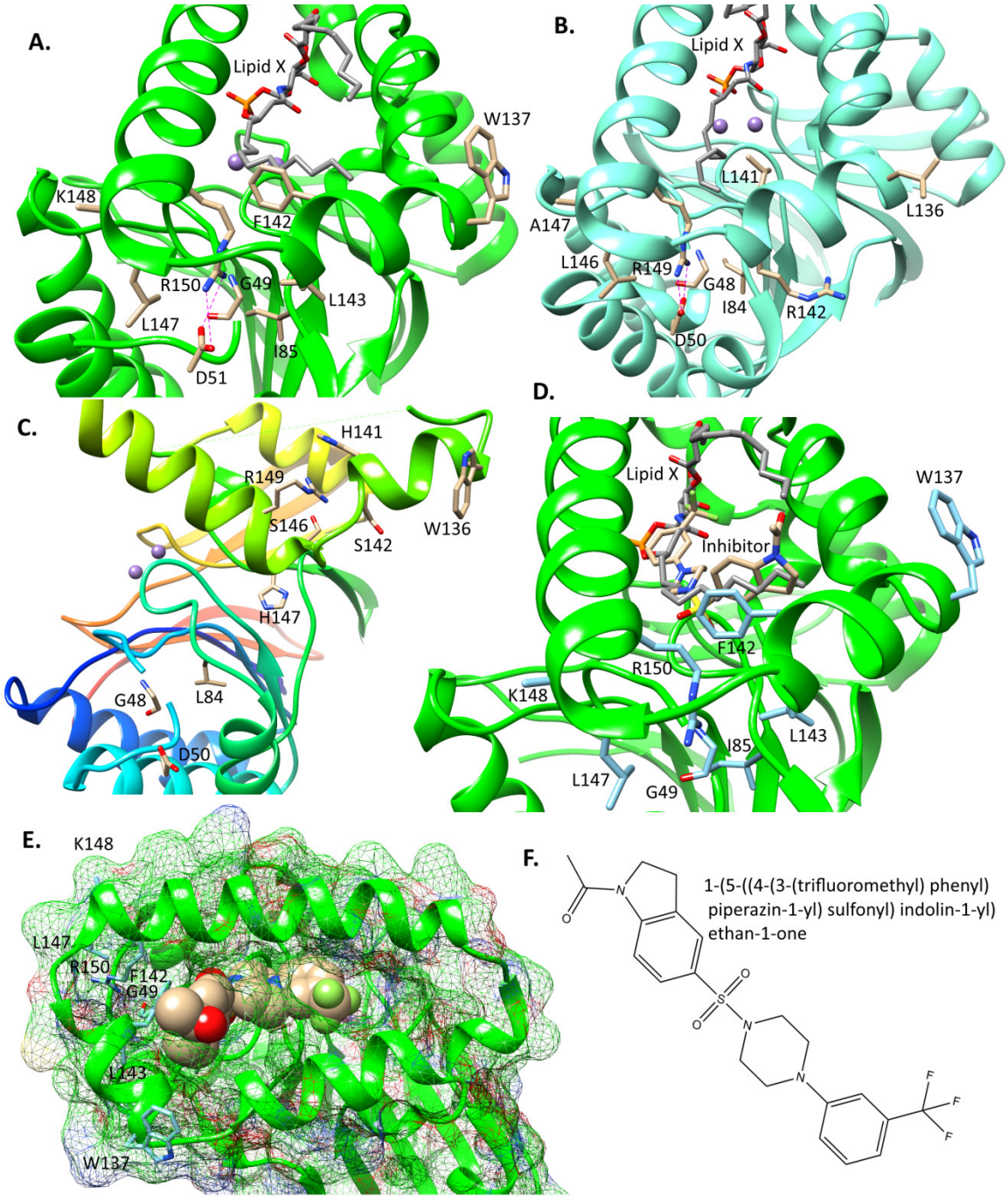


Figure 2.8: Inhibitor binding and cap mobility. Residues that have been shown to be important or that may be important for inhibitor binding are shown. G48D, L84R, F141L, and R149H were found to provide *E. coli* with resistance to the inhibitor by Nayar *et al.* (156). In this study, LpxH4+4, which contains F141H, was also found to be resistant, but reintroduction of F141 in LpxH3+4W136H was not sufficient to allow low micromolar inhibition suggesting some contribution to resistance by one or more of L142S, L146S, and F147H. **A.** HiLpxH (PDB: 5K8K) (154) has residue numbering +1 that of *E. coli* LpxH and

PaLpxH. R150 is forming a ion pair with D51 (3.5 Å). The amide-linked acyl chain of lipid X is interacting with F142. **B.** PaLpxH (PDB: 5B49) (85) has Leu at 141 instead of Phe suggesting that *Pseudomonas aeruginosa* has a natural resistance to the inhibitor. Furthermore, L141 appears to allow the amide-linked acyl chain to assume a conformation different from that observed in HiLpxH. R149 is forming an ion pair with D50 (2.8 and 2.9 Å) **C.** The positions of these residues in the LpxH structure solved in this study. Displacement of the cap has broken the ion pair between R149 and D50. **D.** Docking of the inhibitor (tan carbons) into the HiLpxH structure with AutoDock Vina (157) shows that it may occupy the same space as the amide-linked acyl chain of lipid X (grey carbons). In the docking model selected (4^{th} ranked of 9), the inhibitor is engaged in a π-stacking interaction with F142, which may explain the specificity for Phe over Leu at this position. An overlay of the top 4 docking models is shown in Supplementary Fig 2.8. **E.** Another view of the docking model in **D.** with mesh surface rendering, lipid X removed, and the inhibitor rendered in space-filling model. **F.** The 2D structure of the inhibitor, which was rendered in ChemBioDraw Ultra 14.0.

Finally, we quantified the potency for a previously identified LpxH inhibitor and modeled a plausible binding mode consistent with identified resistance mutations (156). The greater understanding of the structure and dynamics of the LpxH substrate-binding pocket provided by this research will aid in the development of more effective inhibitors and antibiotic leads targeting this step in the synthesis of Gram-negative endotoxin.

However, a few questions remain about the enzymology of LpxH: mainly, it has not been determined how LpxH binds the UMP portion of UDP-DAG or whether the bridging or adjacent active site water attacks the α-phosphate (85, 155). Structures of related Ser/Thr protein phosphatases with phosphate or sulfate bound in the active site suggested that the bridging water is the nucleophile, but structures of LpxH with lipid X bound were consistent with attack by either water (85, 154, 155). These questions could both be answered by the structure of LpxH bound to a non-hydrolyzable analogue of UDP-DAG. In addition, while our activity assays were consistent with collapse of the cap into a closed state in the F82G/L83G mutant, experimental data to directly corroborate the change in cap conformation predicted by molecular dynamics simulations are lacking.

A crystal structure of this mutant might confirm this change in conformation, and further hydrogen-deuterium exchange experiments could reveal any corresponding change in solvent accessibility. Moreover, protein NMR or electron paramagnetic resonance with site-directed labeling by nitroxide-containing spin labels could be used to explore changes in the structure and dynamics of the cap (158, 159). Finally, a crystal structure with inhibitor (**3**) (Scheme 2.1) bound would provide a much higher confidence model for the mechanism of inhibition than docking models, which would be more useful for development of antibiotic lead compounds based on (**3**).

2.4 Experimental Procedures

Cloning, Site Directed Mutagenesis, and Expression

E. coli lpxH was PCR-amplified from DH5α cells (NEB) and inserted into the NdeI (NEB) and XhoI (NEB) sites of pET24a(+) (Novagen). Expression of wild-type LpxH was poor. To improve expression, LpxH was solubilized by the mutation of predicted surface-exposed hydrophobic residues. These residues were predicted from a homology model generated by the Phyre2 server with a distantly related hydrolase, Rv0805 (PDB: 3IB7, 17% identity) (160, 161). Several mutations were tested, and F141H, L142S, L146S, and F147H were identified as the best mutations for improving solubility and expression. However, this protein failed to crystallize. To improve crystallization, surface entropy reduction mutations (E14A, E15A, K161T, and E162A) were introduced, which were identified by the SER server (162). This 8 residue mutant (LpxH4+4) was used for the remainder of experiments except where noted otherwise. A second 8 residue mutant (LpxH3+4-W136H) was also generated in which W136 was

mutated to His instead of F141 to test the effect of the F141H mutation on inhibitor binding as discussed below. In addition, I47C and R149C were added to the LpxH4+4 mutant (LpxH4+4-CC) in an attempt to introduce a disulfide bond that could lock down the capping domain; however, only the presence of intermolecular disulfides was observed by SDS-PAGE (data not shown). Finally, F82G and L83G were added to LpxH4+4 (LpxH4+4-GG) to test the effect on activity as suggested by molecular dynamics simulations (discussed below).

LpxH was expressed in BL21 DE3 cells (Lucigen) grown in Miller's Luria Broth (RPI) (40 mg/L kanamycin [Teknova]) at 37^{o}C to an OD_{600} of 0.8 and induced at 18^{o}C with 1 mM isopropyl β-D-1-thiogalactopyranoside (GoldBio) overnight. The cells were pelleted at 4,500xg (4^{o}C, 25 min) (Beckman J6-MI, JS-4.2), resuspended in lysis buffer (0.5 M NaCl [Fisher], 50 mM Tris-HCl pH 7.4 [Fisher], 5 mM β-mercaptoethanol [β-ME] [EMD Millipore], and 1 mM $MnCl_2$ [MP Biomedicals]), and frozen at -20^{o}C until purification.

Purification of Soluble LpxH

The cells were thawed in a cool water bath and sonicated on ice with a Branson Sonifier (5 output, 50% duty) for three 2 min intervals. The lysate was centrifuged for 45 min at 64,000xg (4^{o}C, 45 min) (Beckman Avanti J-25 I, JA-25.50). The lysate was batch-bound for 1 h (4^{o}C) with 6 mL pre-equilibrated HisPur Ni-NTA resin (Thermo) with the addition of 10 mM imidazole (Chem-Impex Int'l). A gravity-flow column was used to collect the resin, which was then washed with 50 mL equilibration buffer (0.5 M NaCl, 50 mM Tris-HCl pH 7.4, and 5 mM β-ME) with 25 mM imidazole and 15 mL equilibration buffer with 40 mM imidazole. LpxH was eluted in 20 mL equilibration

buffer with 250 mM imidazole. Elutant was dialyzed against 500 mL 0.5 M NaCl, 20 mM Tris-HCl pH 7.4, 5 mM dithiothreitol (DTT) (GoldBio), and 5 mM $MnCl_2$ at 4°C in 3.5 kDa cutoff tubing (Spectra/Por) overnight. Dialyzed elutant was concentrated 4-fold in a 10 kDa cutoff centrifugal filter (Millipore), and concentrated protein was run on HiLoad 26/60 Seperdex 200 column (GE Healthcare) in 0.5 M NaCl, 20 mM Tris-HCl pH 7.4, 5 mM DTT, and 2.5 mM $MnCl_2$ at 4°C. Peak fractions corresponding to monomeric LpxH were concentrated to 14.2 mg/mL, as measured by A_{280} (ε= 28085 $M^{-1}cm^{-1}$) (Nanodrop 8000 Thermo), flash frozen in liquid N_2, and stored at -80°C. The selenomethionine (SeM) derivative of LpxH was produced by the Met synthesis inhibition method with 50-75 mg/L SeM (Chem-Impex Int'l) (163). The SeM derivative was purified as for the native protein except the concentrations of DTT and β-ME were increased to 10 mM.

Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) was utilized to identify additives that would stabilize LpxH for crystallization. Thermal shift assays were performed in a CFX96 Real-Time System C1000 Touch Thermal Cycler (Bio-Rad). Samples (25 µL) were prepared in a 96-well plate with a final concentration of 0.484 mg/mL protein, 80x Sypro Orange (Life Technologies) (1.6% dimethyl sulfoxide [DMSO]), and 2-fold diluted Solubility and Stability Screen (HR2-072 Hampton Research). The temperature was increased from 20°C to 95°C at 0.5°C per cycle with a 30 s cycle. Sucrose (1 M) and 10 mM GSH/GSSG glutathione were identified as strong stabilizing agents.

DSF was also utilized to identify low salt conditions appropriate for hydrogen-deuterium exchange mass spectrometry of LpxH4+4 and to compare the stability of LpxH4+4 and

LpxH4+4-F82G/L83G, which was identified by molecular dynamics simulations as discussed below, in size-exclusion buffer. These experiments were performed as above except that the final concentration of LpxH was 5 μM and sample volume was 40 μL.

Crystallization and X-ray Diffraction

LpxH4+4 (7.1 mg/mL) with 20 mM reduced glutathione (Calbiochem) was crystallized in 1.5 μL hanging drops (2:1 protein to well solution) over 0.1 M Tris-HCl pH 8.2, 70-80 mM magnesium formate (Fluka), and 1-5% 2-propanol (Fisher) at 19^{o}C. Plate-shaped crystals grew within 2 days. Crystals were cryo-protected with 0.25 M NaCl, 10 mM Tris-HCl pH 7.4, 2.5 mM DTT, 1.25 mM $MnCl_2$, 50 mM Tris-HCl pH 8.2, 0.5% 2-propanol, 35 mM magnesium formate, 10 mM reduced glutathione, and 30% 2-methyl-2,4-pentanediol (ACROS Organics). The crystals were shot at the Advanced Photon Source (Argonne National Laboratory, Illinois) on the 24-ID-E beamline at 0.979 Å and 100 K. The best crystals diffracted to 2.0 Å (Table 2.1). The SeM derivative was crystallized under the same conditions as the native protein with the exception of the addition of 5 mM tris(2-carboxyethyl)phosphine-NaOH (Soltec Ventures) pH 6.6 to the well solution. The SeM derivative formed plate-shaped crystals, which were cryo-protected with the same solution as the native crystals. The best SeM derivative crystal diffracted to 2.63 Å at the APS on beamline 24-ID-E at 0.979 Å and 100 K (Table 2.1). The diffraction data were indexed, integrated, and scaled in the beamline's Rapid Automated Processing of Data software (XDS(164) and CCP4(165)).

Model Building

The structure of *E. coli* LpxH was solved in the $P2_122_1$ space group by molecular replacement in PHENIX Phaser with a structure of *P. aeruginosa* LpxH (PDB: 5B4A)

(47% identity), which contained a mutation that eliminated Mn^{2+} binding but was still bound to lipid X, as the search model (85). One hundred and ninety-eight amino acids and one Mn^{2+} ion were built automatically by PHENIX (166). The rest of the structure was built manually in Coot with refinement and automated ligand placement performed in PHENIX (166, 167). Residues 122-130, 162-172, and the C-terminal His-tag are disordered and not visible in the electron density map. The structure was refined with Ramachandran statistics of 94.86% favored, 4.67% allowed, and 0.47% outliers. The SeM derivative data were used to create an anomalous difference map that confirmed the positions of the Se atoms in the structure (Supplementary Fig. 2.4).

Hydrogen-Deuterium Exchange Mass Spectrometry

LpxH was prepared at 20 μM for hydrogen-deuterium exchange in 0.1 M NaCl and 10 mM Tris-HCl pH 8.2. Protein was diluted 10-fold with D_2O at 10^{o}C (100 μL final). Exchange was quenched with an equal volume of quench buffer (1.25% formic acid, 1.5 M guanidine-HCl) at 0 s, 10 s, 1 min, 3 min, 5 min, 10min, 20 min, 30 min, 60 min, and 120 min, and 100 μL was injected on a Waters ACQUITY HDX LC system. The LC system was maintained at 0°C except for a separate chamber for the pepsin digestion (10°C). The samples were digested on a Prozyme pepsin cartridge (2.1 mm x 3.0 mm, Applied Biosystems), trapped, and desalted on a Waters ACQUITY BEH C18 VanGuard Pre-column (130 Å, 1.7 μm, 2.1 mm x 5 mm) for 3 min with a flow rate of 100 μL/min Buffer A (0.1 % formic acid in water). The resulting peptic peptides were resolved on a Waters ACQUITY UPLC BEH C18 column (1.7 μm, 1.0 x 100 mm) with a linear gradient ranging from 5 to 60% acetonitrile with 0.1% formic acid in 6 min at 40 μL/min, which was coupled to a Waters Synapt G2 HDMS q-TOF mass spectrometer.

Figure 2.9: Coverage of the mass spectrometry. Continuous blue bars represent the peptides detected. The maximum coverage was 9 different peptides around residue 200. Gaps show that residues 1-4, 37-61, and 138-140 were not present in any of the observed peptides. The overall coverage was 87.1% with 61 peptides.

Mass spectra were collected in positive ion MS^E mode. Non-deuterated peptides were identified in PLGS 3.0 (Waters), and deuterated peptides were identified and quantified in DynamX 3.0 (Waters) with correction for deuterium incorporation during proteolysis as previously described (168). Hydrogen exchange was measured on 61 peptic peptides, which cover 87.1% primary sequence of LpxH with an average redundancy of 3.51 (Figure 2.9).

Molecular Dynamics Simulations

We performed triplicate molecular dynamics (MD) simulations for four systems: LpxH with lipid X in the binding site ($LpxH_{holo}$), apo LpxH starting from the holo crystal structure ($LpxH_{apo\_holo}$), LpxH F82G/L83G mutant ($LpxH_{mut}$), and apo LpxH starting from the apo crystal structure ($LpxH_{apo\_apo}$). The crystal structures of LpxH from *Pseudomonas aeruginosa* we used as starting structures are PDB: 5B49 (with bound lipid X) for the first three systems and PDB: 5B4C (with no ligand bound) for the last system (85). We used Gaussian09 (169) to generate RESP HF 6-31G* charges and the Antechamber module of the Amber14 suite (170) to parameterize Lipid X with the Generalized Amber Force Field (171). The systems were built with Amber ff14SB force field (172) in a TIP3P (173) water box. Chloride ions were added to neutralize the positively charged systems. Histidine protonation states were determined by the PROPKA program (174) implemented in the NBCR PDB2PQR webserver (175).
All MD simulations were performed using the Amber GPU workflow (176). We first relaxed each system with multi-step minimizations followed by four consecutive restrained MD simulations. The first 2,000 minimization steps constrained all heavy atoms. The second 2,000 minimization steps constrained only the protein heavy atoms and ions. The third 2,000-step minimization was similar to the second minimization but with restraints on the Mn^{2+} ions removed. The fourth 10,000-step minimization held only the protein backbone atoms constrained. The final 20,000 steps minimized all atoms without any constraints. After minimizations, the systems were slowly heated from 0 K to 310 K while keeping a positional restraint of all heavy atoms for 250 ps. Then the systems were subjected to three consecutive 250 ps restrained MD simulations

consecutively releasing the initial 3 kcal/(mol. $Å^2$) restraint on heavy atoms to 2 and 1 kcal/(mol. $Å^2$). Finally, three independent copies of 1 μs production MD simulations with 2 fs time steps were performed at 310 K and 1 atm for all four systems, resulting in 12 μs of MD simulations in total. The MD simulations were stable as shown in root-mean-square-deviation (RMSD) plots (Supplementary Fig. 2.5).

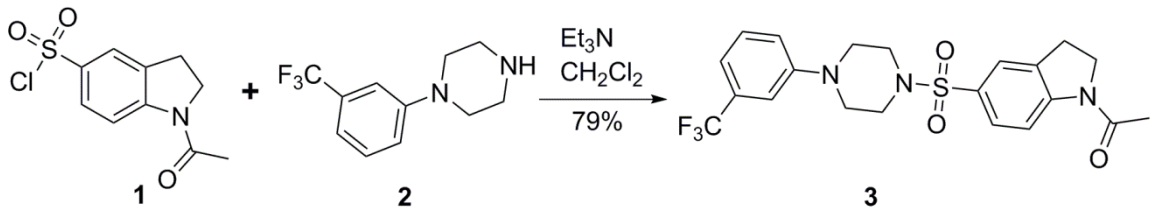
PCA and Distance Analysis

Principle component analysis was performed on all the systems using Amber cpptraj (177). The motion of the cap (residues 121-168) and an interacting loop (residues 80-83) was captured in a covariance matrix, diagonalized to find the eigenvalues, and the conformation variation of each system was shown by projecting onto the eigenvalues. Distances were calculated using VMD (178). R was used for plotting (179).

Chemical Synthesis of LpxH Inhibitor

General Procedures: All commercial chemicals were used as supplied unless otherwise indicated. Flash chromatography was performed on a Teledyne Combiflash RF-200 with RediSep columns (silica) and indicated mobile phase. All moisture sensitive reactions were performed under an inert atmosphere of ultrapure argon with oven-dried glassware. ^1H was recorded on a Varian 600 MHz. Mass data was acquired on an Agilent 1100 LC/MSD TOF mass spectrometer (G1969A). Analysis of sample purity was performed on a Varian Prepstar SD-1 HPLC system with a Phenomenex Gemini, 5 μm C18 column (250 mm × 4.6 mm). HPLC conditions: solvent A = H_2O, solvent B = acetonitrile; flow rate = 1.0 mL/min; compounds were eluted with a gradient of 5% acetonitrile/H_2O for 0−5 min and to 95% acetonitrile/H_2O from 5 to 30 min followed by

100% acetonitrile from 35 to 40 min. Purity was determined by total absorbance at 254 nm. Compound has a purity ≥95%.



Scheme 2.1: Synthesis of LpxH inhibitor. 1-(5-((4-(3-(trifluoromethyl) phenyl) piperazin-1-yl) sulfonyl) indolin-1-yl) ethan-1-one (**3**) was synthesized from piperazine derivative (**2**), 1-acetylindoline-5-sulfonyl chloride (**1**), and triethylamine dissolved in dichloromethane. The reaction was performed at ambient temperature for 2 h with stirring. Following purification, the final yield was 79% at ≥95% purity.

Synthesis of 1-(5-((4-(3-(trifluoromethyl) phenyl) piperazin-1-yl) sulfonyl) indolin-1-yl) ethan-1-one (**3**) (Scheme 2.1): To a solution of piperazine derivative (**2**) (100 mg, 0.43 mmol, 1.0 eq) in dichloromethane (5 mL) at 0 °C, 1-Acetylindoline-5-sulfonyl chloride (**1**) (0.12 g, 0.48 mmol, 1.1 eq) and triethylamine (0.08 mL, 0.56 mmol, 1.3 eq) was added dropwise and the reaction mixture was stirred for 2 h at room temperature. The reaction was diluted with dichloromethane (10 mL) and quenched with sodium bicarbonate (10 mL). The organic layer was separated and the aqueous layer was extracted with dichloromethane (2 x 10 mL). The combined organic layers were washed with brine, dried over sodium sulfate and evaporated in vacuo to yield the desired crude product. Purification of the crude using Combiflash with a gradient of 0-2% methanol in dichloromethane yielded the title compound as colorless crystals (0.16 g, 79%).

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.21 (d, J = 8.2 Hz, 1H), 7.59 (d, J = 11.0 Hz, 2H), 7.40 (t, J = 7.9 Hz, 1H), 7.19 (d, J = 8.4 Hz, 1H), 7.15 (s, 1H), 7.09 (d, J = 7.5 Hz, 1H), 4.17 (t, J = 8.4 Hz, 2H), 3.32 (s, 4H), 3.23 (t, J = 8.4 Hz, 2H), 2.99 (s, 4H), 2.19 (s, 3H). HRMS-ESI (+) *m/z* calculated for $C_{21}H_{23}F_3N_3O_3S$, 454.1412 [M+H]-; found: 454.1406.

Purification of LpxH-mCherry

In order to test the inhibition of wild-type LpxH, a C-terminal fusion with mCherry was designed. The *lpxH* and *mCherry* genes were fused together with a KpnI (NEB) site followed by 3 consecutive Gly codons. The gene fusion was ligated into the BamHI (NEB) and XhoI sites of pET24. This fusion allowed for the purification of enough LpxH-mCherry for enzymatic assays. The Ni-NTA portion of the purification process for this protein was the same as for the native soluble form of LpxH except that β-ME was decreased to 1 mM, Tris-HCl pH 7.4 was replaced with HEPES-NaOH pH 8.0 (Fisher), and 10% glycerol (Fisher) was added to the buffer. The elutant (15 mL) was dialyzed against 500 mL 5% glycerol, 0.5 M NaCl, 50 mM HEPES-NaOH pH 8.0, 5 mM $MnCl_2$, and 1 mM DTT at 4^{o}C in 3.5 kDa cutoff tubing overnight. The dialyzed elutant was concentrated in a 30 kDa cutoff centrifugal filter (Millipore), and concentrated protein was run on a Seperdex 200 Increase 10/300 GL column in 5% glycerol, 0.5 M NaCl, 50 mM HEPES-NaOH pH 8.0, 5 mM β-ME, and 2.5 mM $MnCl_2$ at 4^{o}C. The most pure fractions were combined and the concentration was measured as 0.413 mg/mL (7.52 μM) as measured by A_{280} (ε= 62340 $M^{-1}cm^{-1}$).

Activity Assays

For TLC-based assays, LpxH variants were prepared at 100 nM in 0.5 M NaCl, 20 mM Tris pH 7.5, and 2.5 mM $MnCl_2$. Reactions (10 μL) were prepared with 62 μM UDP-DAG (as measured by A_{260}, ε= 9.9 $mM^{-1}cm^{-1}$), 10 nM LpxH or 1 μL dilution buffer, 100 mM Tris-HCl pH 8.0, 0.1% Triton X-100, and 1 mg/mL bovine serum albumin. UDP-DAG and lipid X were prepared as described in Chapter 3. Reactions were run at 30^{o}C and were quenched by spotting onto an HPTLC Silica gel 60 plate (EMD

Millipore). The spots were dried with a Dual-Temp heat gun (Genesis) set on low (300^{o}C). The plate was run with 25/15/4/2 chloroform/methanol/water/acetic acid mobile phase. After running, the plate was allowed to dry for one hour. Then the plate was developed by spraying with 20% sulfuric acid in ethanol and charring with the heat gun set on high (538^{o}C) until bands appeared. Reactions to test the effect of the LpxH inhibitor contained 10% DMSO (TCI) and were run for the amount of time required for the uninhibited enzyme to reach completion (Supplementary Table 2.2) as determined above.

While TLC-based assays have the benefit of directly showing changes in the substrate and product of interest, absolute quantification requires concurrently run standard curves (180, 181). To determine specific activities for more quantitative comparisons of LpxH variants, UMP concentration was quantified with a UMP/CMP-Glo Glycosyltransferase Assay kit (Promega) as described by the manufacturer. Reactions (10 μL) were prepared as above except that enzyme dilution buffer included 5 mM DTT, reaction LpxH concentration was 5 nM (20 nM for LpxH4+4-GG), and UDP-DAG concentration was 0.1 mM. Reactions were run at ambient temperature (21^{o}C) in a white 384-well plate (Greiner Bio-One 781074). Reactions were quenched with 10 μL UMP/CMP Detection Reagent (Promega), and the plate was shaken 30 s at 1440 rpm and 1 mm amplitude in a Spark 10M plate reader (Tecan). The luciferase reaction was incubated at ambient temperature 1 h, and then the total luminescence of each well was measured in a luminometer developed by Fluorescence Innovations (Minneapolis, MN). Concurrently run UMP standard wells allowed luminescence of LpxH reaction wells to be converted to UMP concentration. All reactions and standards were run in triplicate. To

determine specific activities, reactions were quenched at three time points within the early, linear product accumulation range of the LpxH reactions, and linear regression was performed in Graphpad Prism v7.03. To determine the IC_{50} of the LpxH inhibitor, specific activities of LpxH-mCherry (5nM) were determined as above with 10% DMSO and 2-50 μM inhibitor. These specific activities with standard errors were fit to a 3-variable dose-response inhibition curve in Graphpad Prism v7.03: $SA = SA_{min} + \frac{SA_{max} - SA_{min}}{1+10^{\log[inhibitor] - \log(IC50)}}$. For the purpose of curve fitting, an inhibitor concentration of 0.1 nM was paired with the uninhibited (10% DMSO) LpxH-mCherry specific activity.

Molecular Docking

The LpxH inhibitor 1-(5-((4-(3-(trifluoromethyl) phenyl) piperazin-1-yl) sulfonyl) indolin-1-yl) ethan-1-one (**3**) (Scheme 2.1) (156) was docked into the HiLpxH (PDB: 5K8K) structure (154) with AutoDock Vina (157). The protein and ligand structures were prepared for docking in AutoDockTools from the MGLTools 1.5.6 suite (182, 183). The active site was defined by a box centered at coordinates (0.2, 16.529, 13.923) with x, y, z lengths of 24, 18, and 36 Å. Exhaustiveness was set to 10; 9 structures were output.

Data Availability

The coordinates and structure factors of the crystal structure reported in this article are available at the Protein Data Bank under accession number 5WLY. All other data presented are available within this article or its supplementary information or from the corresponding authors by request.

2.5 Supplementary Information

Supplementary Table 2.1: Thermal Stability

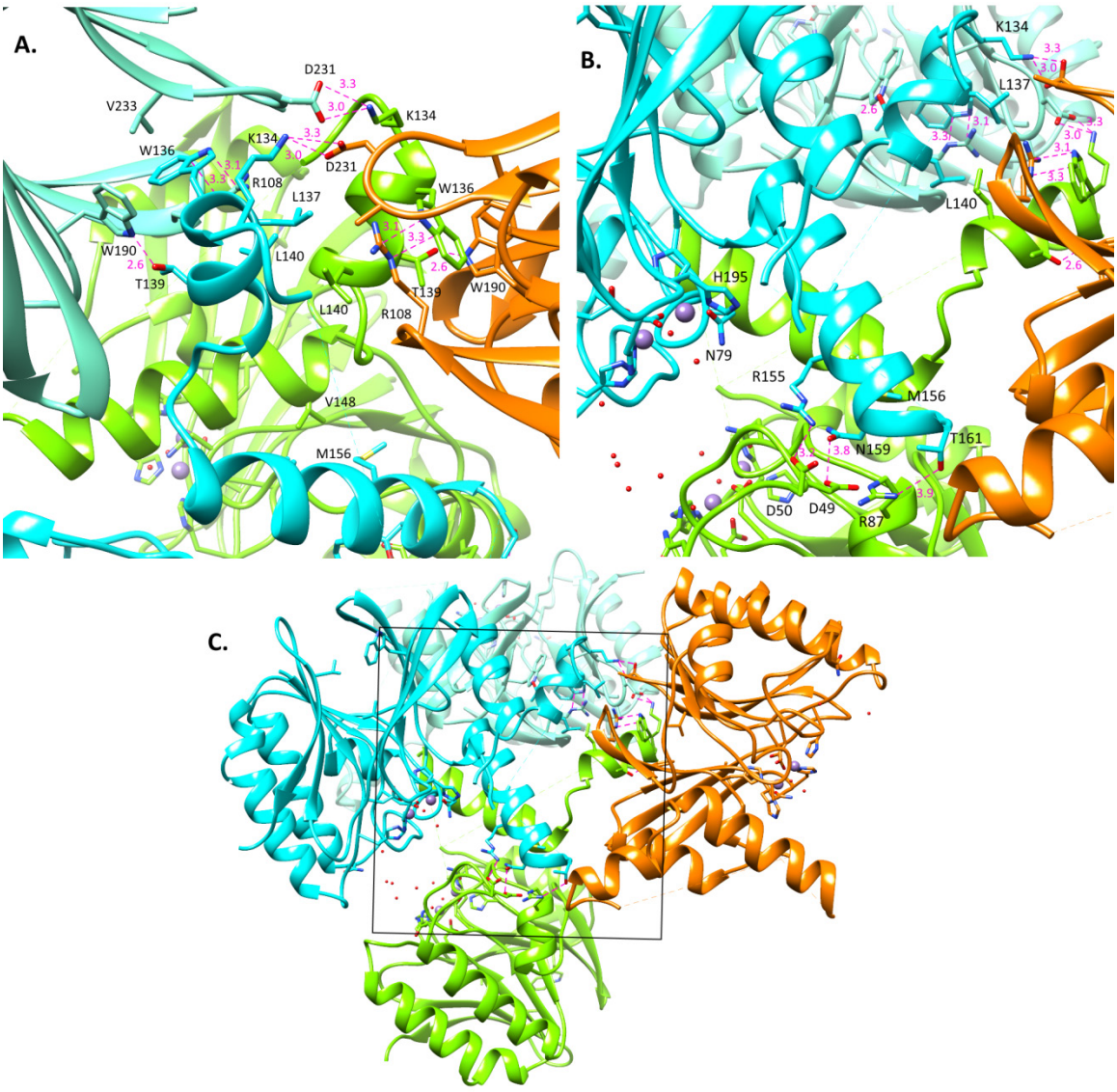
LpxH	Melting Temperature (^{o}C)
4+4	60.3±0.3
4+4-GG	57.7±0.3

LpxH variants and Sypro Orange were prepared at 5 μM and 80x, respectively, in 0.5 M NaCl, 20 mM Tris pH 7.5, 5 mM DTT, and 2.5 mM $MnCl_2$. Fluorescence was measured as temperature was increased stepwise from 20^{o}C to 95^{o}C. Average T_m values with standard deviations are shown (n=3).

Supplementary Table 2.2: TLC-Based Time to Reaction Completion

LpxH	Time to reaction completion
-mCherry	20 min <t< 30 min
4+4	15 min <t< 20 min
3+4-W136H	30 min <t< 40 min
4+4-CC	40 min <t< 50 min
4+4-GG	5 h <t*

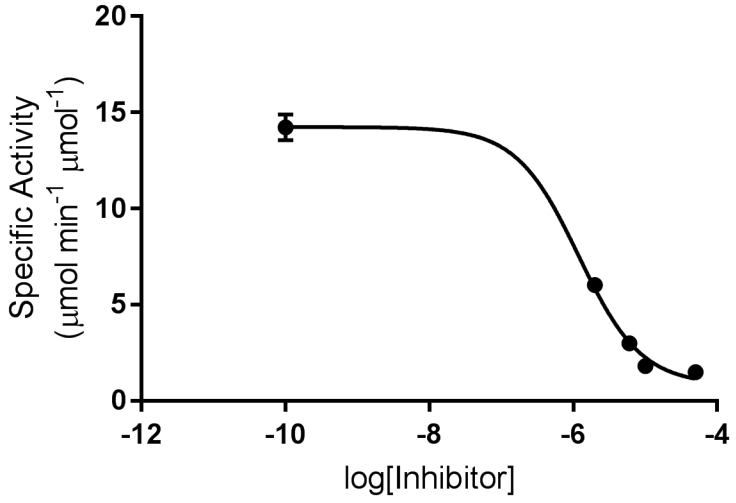
Reactions (10 μL) were run at 30^{o}C with 62 μM UDP-DAG, 10 nM LpxH, 0.1 M Tris-HCl pH 8.0, 0.1% Triton X-100, and 1 mg/mL BSA. *When LpxH concentration was increased to 100 nM, t was between 3 and 4 h.



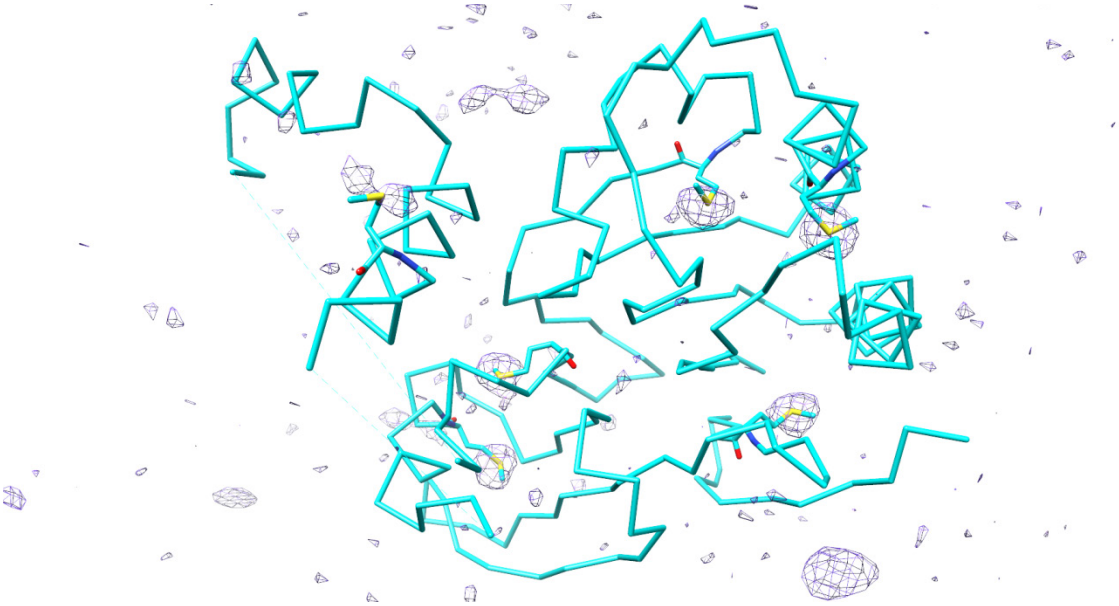
Supplementary Fig. 2.1: Crystal contacts that stabilize helical cap. A. The interactions between the first of the visible capping helices (helix 6) and the adjacent proteins in the crystal. K134 Nζ is interacting with D231 Oδ1 and Oδ2 at 3.3 and 3.0 Å, respectively. W136 is engaged in a cation-π stacking interaction with R108 (3.1 Å between Trp Nε and Arg Cζ and 3.3 Å between Trp Cε2 and Arg Nε). T139 is hydrogen-bonding with W190 (2.6 Å between Thr Oγ and Trp Nε). **B.** The interactions between the second visible helix of the capping domain (helix 7) and the neighboring protein in the crystal. R155 Nη2 is interacting with D50 Oδ1 at 3.2 Å. N159 Oδ is interacting with D49 Oδ1 at 3.8 Å, and T161 Oγ is interacting with R87 Nη at 3.9 Å. **C.** Full view of protein chains. The box indicates the portion shown in **B**.

```
EcLpxH -MATLFIADLHLCVEEPAITAGFLRFLAGEARKADALYILGDLFEAWIGDDDPNPLHRKM
PaLpxH -MSVLFISDLHLEAERPDITRAFLSFLDERARRAEALYILGDFFEAWIGDDGMDAFQRSI
HiLpxH MKHSYFISDLHLSETQPELTALFVDFMQNLAPQAERLYILGDLFDFWIGDDEQSALIQQV
           .** ****. ..*.:*  *:.*:   * .*: ******.*: *****  . . ..:

EcLpxH AAAIKAVSDSGVPCYFIHGNRDFLLGKRFARESGMTLLPEEKVLELYGRRVLIMHGDTLC
PaLpxH AQSLRQVADGGTRIYLMHGNRDFLIGKAFCREAGCTLLPDPSVIDLYGEPVLLMHGDSLC
HiLpxH KDLIKFVSDQGVQCYFQHGNRDFLIGERFSKETGAQLLPDYQLITLYDKKILLCHGDTLC
          :: * * *   *. *********. * :* * .***: .::.** . :*: ***:**

EcLpxH TDDAGYQAFRAKVHKPWLQTLFLALPLFVRKRIAARMRANSKEANSSKSLAIMDVNQNAV
PaLpxH TRDEAYMRLRRWLRNPLTLWVLRHLPLATRHKLARKLRKESRAQTRMKAVDIIDVTPEEV
HiLpxH IDDEAYQQFRRRVHQKWLQRLFLCLPLKVRVIIAEKIRAKSNQDKQAKSQEIMDVNQAFT
         .*  *  .*  :.. .   :.  ***  *  :* ::* .*.  .. *   *:**.

EcLpxH VSAMEKHQVQWLIHGHTHRPAVHELIANQQPAFRVVLGAWHTE-GSMVKVTADDVELIHF
PaLpxH PRVMRGHGVRTLIHGHTHRPAEHPLDIDGQPARRIVLGDWDRQ-GWALEIDANGHRQAPF
HiLpxH AEKVQEFGVNLLIHGHTHREAIHQ----QEEFTRIVLGDWRKNYASILKMDESG----EF
         . :.   *. ******** * *      :   *:*** *..:    :.:. .      *

EcLpxH PF---
PaLpxH PL---
HiLpxH GFIKD
```

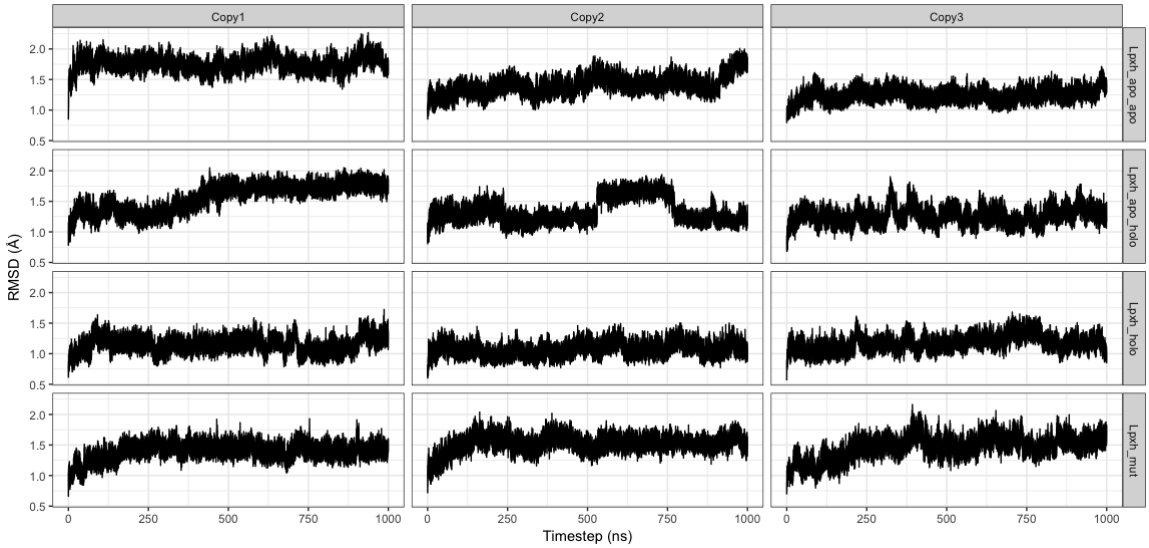
Supplementary Fig. 2.2: LpxH sequence alignment. Protein sequences of *E. coli*, *H. influenzae*, and *P. aeruginosa* LpxH were aligned in MEGA6 by the Muscle algorithm with default parameters (184). The degree of conservation between the sequences is designated by "*" for conserved residues, ":" for similar residues (aliphatic, aromatic, basic, acidic, small hydrophilic, etc.), and "." for residues of the same class (hydrophobic vs. hydrophilic). Yellow highlighting indicates residues mutated in LpxH4+4. Green highlighting indicates other residues mutated in this study. Red lettering corresponds to the helical cap domain. The sequence identity between *E. coli* and *P. aeruginosa* LpxH is 47.1%. The sequence identity between *E. coli* and *H. influenzae* LpxH is 46.1%. The overall sequence identity is 33.5%.



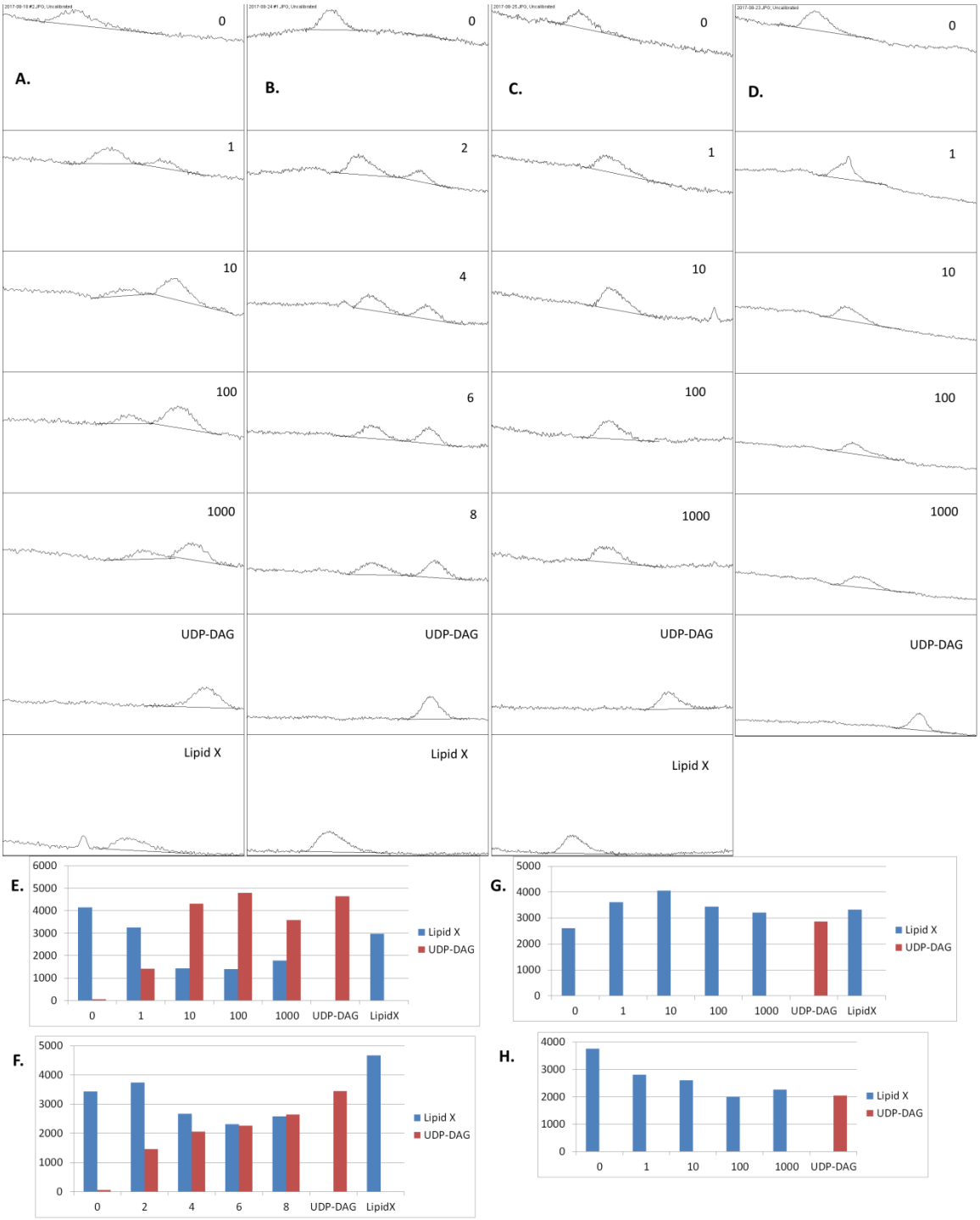
Supplementary Fig. 2.3: LpxH-mCherry inhibition. Specific activities with standard errors shown in Table 2.2 were fit to a 3-variable dose-response inhibition curve in Graphpad Prism v7.03. To allow inclusion of the uninhibited control, the specific activity of the DMSO control was paired with an inhibitor concentration of 0.1 nM. The best fit IC_{50} value of inhibitor **3** (Scheme 2.1) was 1.188±0.2 μM ($Log_{10}[IC_{50}]$= -5.925±0.077 with a 95% confidence interval of 0.4648 to 2.35 μM.



Supplementary Fig. 2.4: Methionine positions. The selenium SAD data was utilized with the refined native model to generate an anomalous difference map. The contour mesh is displayed at 3.00 RMSD (0.0672 e/Å^3) and confirms the location of the methionines built in native model.

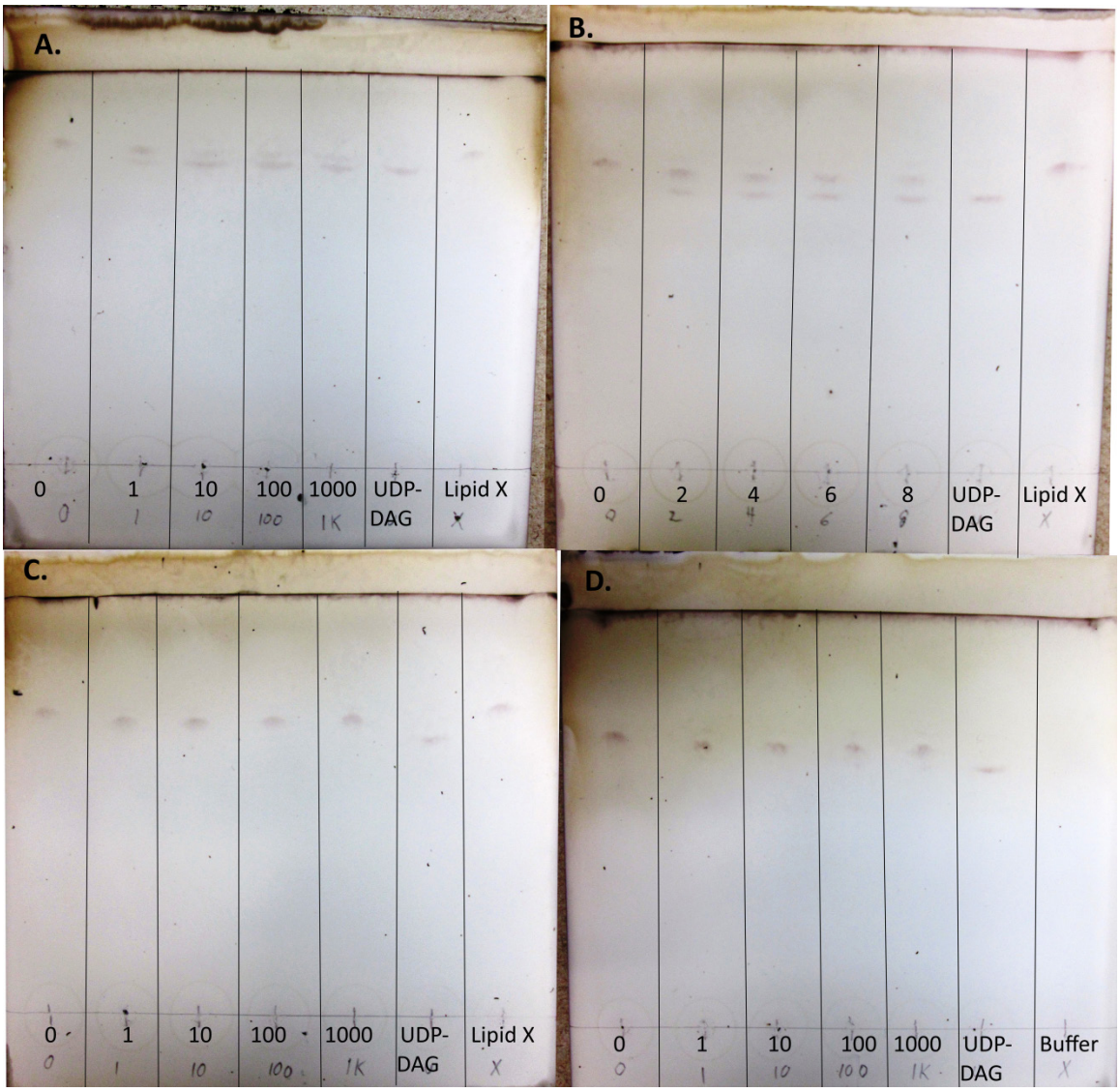


Supplementary Fig. 2.5: Molecular dynamics RMSD plots. RMSD of the simulated systems aligned by Cα atoms.



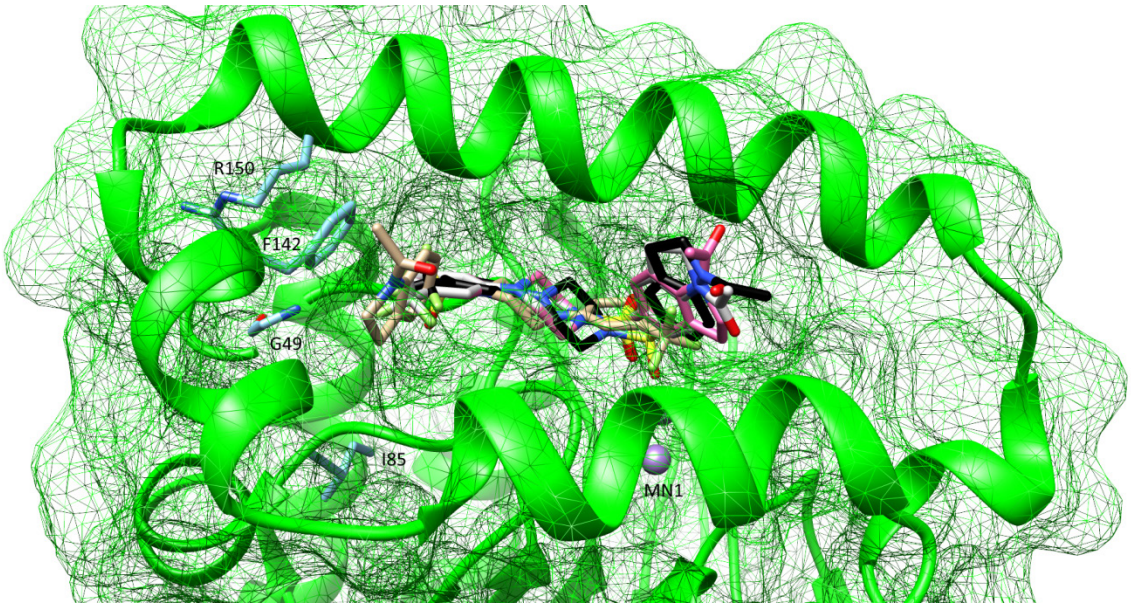
Supplementary Fig. 2.6: Quantification of lipid band intensities. Band intensities from Figure 2.7 were analyzed in ImageJ (185). Lanes containing the entirety of the bands were defined with rectangular selections, and the lanes were plotted as distance versus intensity. The areas under the peaks were determined and graphed with x-axis numbers indicating inhibitor concentration [μM]. **A.** and **E.** and **B.** and

F. correspond with LpxH-mCherry. **C.** and **G.** correspond with LpxH4+4. **D.** and **H.** correspond with LpxH3+4-W136H.



Supplementary Fig. 2.7: Full TLC plates. Reactions (10 µL) were run at 30°C with 62 µM UDP-DAG, 10 nM LpxH, 0.1 M Tris-HCl pH 8.0, 0.1% Triton X-100, 10% DMSO, and 1 mg/mL BSA at various concentrations of the inhibitor **3** (Scheme 2.1) (shown in micromolar). Reactions were quenched by spotting on HPTLC Silica gel 60 plates, which were run with 25/15/4/2 chloroform/methanol/water/acetic acid mobile phase and charred with 20% sulfuric acid in ethanol and a heat gun. Reactions prepared with UDP-DAG but no enzyme act as a negative control and UDP-DAG standard. Independently purified lipid X product is also included as a standard. **A.** and **B.** Plates for LpxH-mCherry. These reactions were run for 30 min. The reaction proceeded to completion in the absence of inhibitor, but is slightly inhibited by 1 µM inhibitor. At 2-8 µM inhibitor, the reaction proceeds to about 50% completion, and it is mostly inhibited at 10 µM and above. These plates were not repeated. **C.** LpxH4+4 reactions were run for 20 min. No

inhibition was observed up to 1 mM inhibitor. This plate was replicated twice. **D.** LpxH3+4-W136H reactions were run for 40 min. A small amount of UDP-DAG remains at 100 μM and 1 mM inhibitor suggesting weak inhibition at high inhibitor concentrations. This plate was replicated thrice. Analysis of band intensities is presented in Supplementary Fig. 2.6.



Supplementary Fig. 2.8: Top ranked docking poses. The top 4 docking poses of inhibitor **3** (Scheme 2.1) predicted by AutoDock Vina (157) are overlaid with color-coding of carbon atoms: 1 pink, 2 grey, 3 black, and 4 tan. In model 4 (also shown in Figure 2.8), the molecule is rotated ~180^{o} relative to the previous poses, allowing a putative π-stacking interaction between the bicyclic ring and F142.

CHAPTER 3: CRYSTAL STRUCTURE OF THE LIPID A DISACCHARIDE SYNTHASE LpxB FROM *Escherichia coli*

3.1 Background

LpxB is a glycosyltransferase in the Raetz (lipid A synthesis) pathway that catalyzes nucleophilic attack of the 6-hydroxyl of lipid X on the anomeric carbon of UDP-diacyl-glucosamine (UDP-DAG) to form β(1-6)-tetraacyl-disaccharide 1-phosphate (lipid A disaccharide) (Figure 3.1) (39–41, 47, 186).

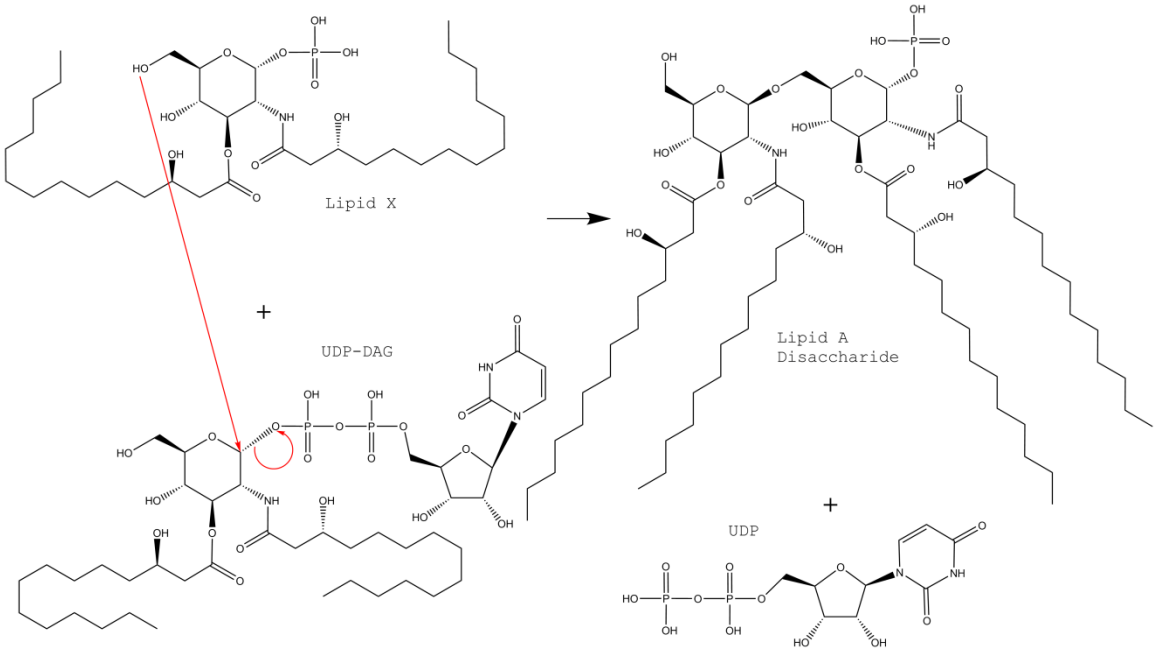


Figure 3.1: LpxB reaction. LpxB catalyzes the nucleophilic attack of the 6'-hydroxyl of lipid X on the anomeric carbon of UDP-DAG with UDP as the leaving group (40). As for other inverting GT-B enzymes, this reaction is thought to proceed by an SN2 mechanism (98, 187).

LpxB is essential for growth of *E. coli* and is among the most highly conserved enzymes in the Raetz pathway (7, 188). *LpxB* was first identified from a temperature sensitive mutant of *E. coli* (*pgsB1*) that was shown to accumulate lipid X and lipid Y (a palmitoylated form of lipid X) at 42°C (45, 46, 189). An *lpxB* mutant was also found to accumulate UDP-DAG, and characterization of the activity and products of LpxB brought about a breakthrough in the understanding of lipid A structure (40, 153). LpxB was fully purified and enzymatically characterized by Metzger and Raetz who showed

that LpxB is a membrane surface active enzyme (186). Despite being the founding member of the Raetz pathway, LpxB remained structurally uncharacterized until the present research (40).

Of the enzymes present in the canonical Raetz pathway, 7 have been structurally characterized (LpxA (62), LpxC (72), LpxD (80), LpxH (85, 154), LpxK (91), WaaA (96), and LpxM (107)) and 2 have not (LpxB and LpxL). In addition, LpxI (a nonhomologous alternative to LpxH) has been structurally characterized by x-ray crystallography (82, 87). Antibiotic lead compounds that target LpxC (190–196), LpxH (156), LpxD (197), and LpxA (197, 198) have been developed, but none have been reported for LpxB. This article presents the first crystal structures of a soluble variant of *E. coli* LpxB in the apo form and bound to UDP, which reveals a novel glycosyltransferase-B (GT-B) dimer wherein the C-terminal tail of one subunit completes the fold of the other subunit. We further demonstrate the importance of LpxB dimerization via C-terminal swapping through functional analyses.

3.2 Results

Overall Structure of LpxB

In order to crystallize LpxB, a soluble form of the enzyme was generated. We utilized the Phyre2 server (160) to create a homology model of LpxB based on UDP-N-acetylglucosamine 2-epimerase from *Thermus thermophilus* HB8 (PDB: 1V4V). We compared the homology model to the structure of MurG (PDB: 1F0K and 1NLM) (97, 101), which was found to have a hydrophobic patch likely involved in membrane association (97).

Table 3.1: Diffraction and Refinement Statistics

Parameter	LpxB7S (PDB: 5W8S)	LpxB7S SeMet (SAD data)	LpxB7S+UDP (PDB: 5W8X)	LpxB6S (PDB: 5W8N)
Diffraction				
Wavelength (Å)	0.97918	0.97912	0.9792	0.97919
Resolution range (Å)	155.14-2.10 (2.16-2.10)	96.29-3.43 (3.71-3.43)	153.88-1.98 (2.03-1.98)	19.57-2.02 (2.07-2.02)
Space group	P 32 2 1	P 64	P 32 2 1	P 32 2 1
Unit cell a,b,c (Å) α,β,γ (°)	68.02 68.02 155.1 90 90 120	111.2 111.2 74.19 90 90 120	67.81 67.81 153.9 90 90 120	67.78 67.78 154.50 90 90 120
Total observations	87521 (7195)	143695 (30114)	118812 (7310)	180372 (11983)
Total unique	24865 (2012)	7129 (1469)	29131 (1983)	27746 (1958)
Multiplicity	3.5 (3.6)	10.4 (10.4)	4.1 (3.7)	6.5 (6.1)
Completeness (%)	99.4 (99.9)	100 (100)	99.1 (97.8)	99.6 (97.0)
Mean I/σI	13.5 (1.6)	21.6 (1.9)	13.2 (1)	14.2 (1.2)
Wilson B-factor	40.03	33.07	35.15	38.57
R-merge	0.066 (0.997)	0.097 (1.926)	0.063 (1.272)	0.112 (1.686)
R-meas	0.078 (1.176)	0.102 (2.025)	0.073 (1.485)	0.122 (1.838)
R-pim	0.041 (0.615)	0.032 (0.623)	0.035 (0.753)	0.047 (0.717)
CC1/2	0.998 (0.55)	0.999 (0.732)	0.999 (0.45)	0.997 (0.530)
Refinement		Not fully refined		
Resolution Range (Å)	55.07-2.10 (2.15-2.10)		54.86-1.98 (2.05-1.98)	19.57-2.02 (2.09-2.02)
Completeness (%)	98.97 (99)		90.65 (75)	99.62 (98)
Reflections	24788 (1737)		26651 (2153)	27680 (2650)
R-free Reflections	2007 (141)		1373 (102)	1385 (131)
R-work/R-free	0.1951/0.2226 (0.3344/0.3789)		0.1990/0.2309 (0.3512/0.3278)	0.1872/0.2129 (0.3043/0.3439)
Non-hydrogen atoms	2953		2964	3043
Protein	2800		2817	2859
Ligands/ions	3		25	0
Solvent	150		122	184
RMS bond lengths (Å)	0.003		0.003	0.005
RMS bond angles (°)	0.507		0.572	0.623
Ramachandran favored (%)	98.63		98.34	99.18
Ramachandran allowed (%)	1.10		1.10	0.55
Ramachandran outliers (%)	0.27		0.55	0.27
Rotamer outliers (%)	0.00		0.69	1.01
C β -outliers	0		0	0
Clashscore	0.71		1.76	0.70
Average B-factor	45.6		46.12	46.06
Protein	45.39		46.06	45.80
Ligands/ions	51.25		52.67	
Solvent	49.44		46.15	50.20
Number of TLS groups	1		5	4

Statistics for highest resolution shell are shown in parentheses. Each dataset corresponds to one crystal.

We identified, and mutated to Ser, 6 Val and Leu residues (V66, V68, L69, L72, L75, and L76) in a nearby region of the LpxB homology model. These mutations improved solubility and yield of LpxB, and the LpxB with all 6 mutations (LpxB6S) showed the least aggregation on a size exclusion column (Supplementary Fig. 3.1). LpxB6S was used for initial crystallization screening and optimization. An additional mutation (M207S) was added for selenomethionine derivative crystallization giving LpxB7S (Table 3.1).

LpxB6S was crystallized with one molecule in the asymmetric unit. However, the biological assembly is a highly intertwined, C-terminally swapped homodimer. Each half of the LpxB dimer consists of two Rossmann-like domains with a parallel β-sheet core sandwiched between α-helices (Figure 3.2 and Supplementary Fig. 3.2). Both polypeptides within the LpxB dimer contribute to the fold of each domain. The first Rossmann-like domain is composed of the N-terminal half of one subunit [residues 1-166] and the last helix [369-382] from the C-terminus of the other subunit. This domain is connected via short antiparallel helical linkers to the second domain, which is composed of residues 177-295 of the first subunit and 296-345 of the second subunit. While the tertiary structure of LpxB is typical for the glycosyltransferase-B (GT-B) family, the quaternary structure is unique (98, 187). LpxB forms a two-fold symmetric dimer in which the polypeptide chains crossover at helix 14 around the dyad axis to swap the remainder of the polypeptide chain. The reciprocally swapped C-terminal segments wrap around the opposing subunit in an interlocked arrangement. These extensive interactions bury a total surface area of 12,650 $Å^2$, forming an LpxB dimer with an overall “bowtie”-shape. Supplementary Fig. 3.2 presents the isolated monomer and a 2D depiction (199) of the dimer.

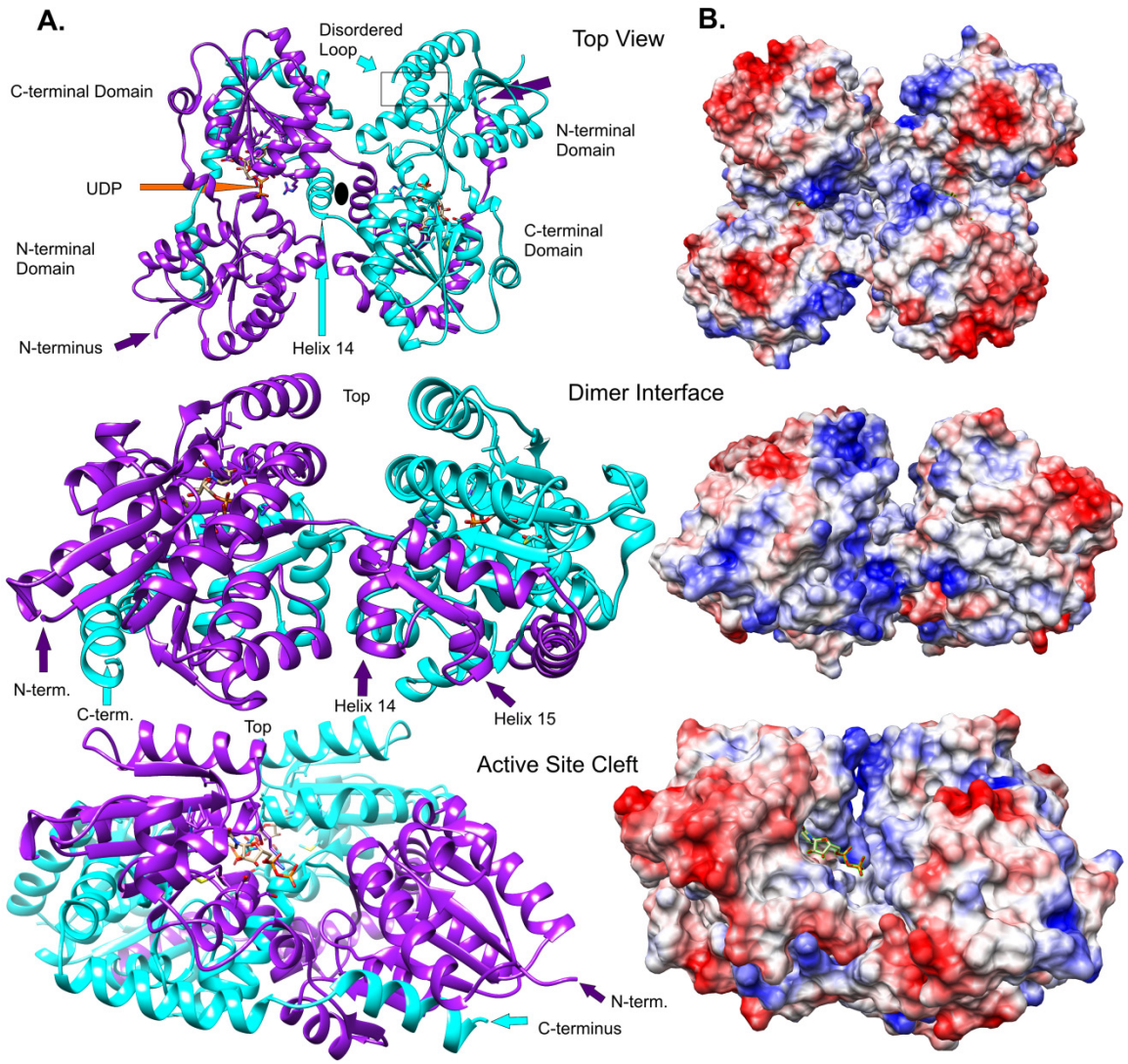


Figure 3.2: Overall structure of LpxB. A. The overall structure of LpxB7S shows that LpxB forms a dimer in which the C-terminus of one subunit completes the fold of the opposite subunit. The polypeptides are exchanged at helix 14. Each domain forms a Rossmann-like fold with a parallel β-sheet surrounded by α-helices. UDP is bound to the C-terminal domain showing the position of the UDP-binding pocket where the sugar donor substrate, UDP-DAG, binds. Lipid X likely binds to the N-terminal domain which includes the predicted catalytic base D98 (186). LpxB appears to be in an open conformation. Catalysis may involve partial closing of the cleft between the domains by a hinge-like movement of the N-terminal domain similar to conformational changes observed for MurG (101), PimA (102), and other GT-B enzymes (98). **B.** Columbic surface rendering (blue for positive and red for negative) corresponding to the adjacent ribbons in **A**. The surface near the dimer interface is highly basic and is likely involved in membrane association.

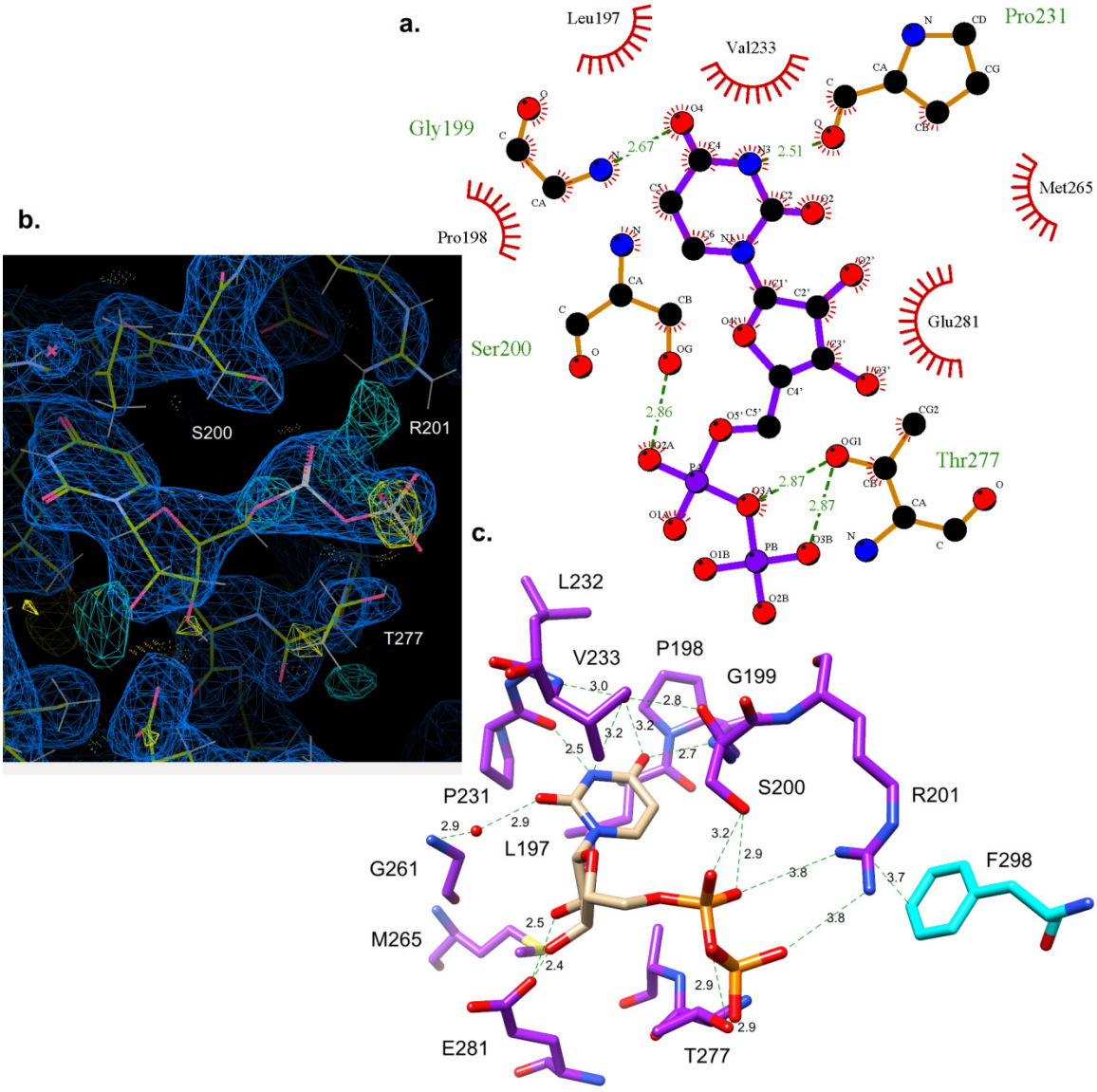


Figure 3.3: LpxB nucleotide-binding pocket. The uracil of UDP binds in a hydrophobic pocket formed by L197, P198, P231, and V233. Putative hydrogen-bond distances in Ångstroms are as follows: G199 N to uracil O4 (2.7), G199 O to water to uracil O4 (2.8, 3.2), V233 N to water to uracil O4 (3.0, 3.2), P231 O to uracil N3 (2.5), G261 N to water to uracil O2 (2.9, 2.9), E281 Oε2 to ribose O2' (2.5), E281 Oε2 to ribose O3' (2.4), S200 Oγ to α-phosphate O1 (3.2), S200 Oγ to α-phosphate O2 (2.9), T277 Oγ to bridging pyrophosphate O (2.9), and T277 Oγ to β-phosphate O2 (2.9). R201, which is critical for activity (186), is interacting with the α-phosphate O2 and β-phosphate O1 at 3.8, via its η-nitrogens and appears to be positioned to stabilize the UDP leaving-group. **A.** Image generated in LigPlot+ (78). **B.** Image generated in Coot (167), showing the electron density maps for UDP and the surrounding residues. The 2mFo-DFc map (blue mesh) is shown at 0.4317 e/Å^3 (0.99 RMSD), and the mFo-DFc map (cyan mesh for positive and

yellow mesh for negative) is shown at 0.45 e/Å^3 (3.01 RMSD). See Supplementary Fig. 3.6 for stereo view. **C.** Image generated in UCSF Chimera (25).

UDP Binding Site

The ligand-bound structure of LpxB was obtained by soaking crystals with 10 mM UDP-N-acetylglucosamine (UDP-GlcNAc), a mimic of one of the reactants, UDP-DAG (Figure 3.3). Only the UDP portion of the molecule is resolved in the electron density map (Figure 3.3B) suggesting that the GlcNAc moiety is highly flexible or that UDP-GlcNAc was hydrolyzed during soaking as has been observed before for nucleotide-charged sugars soaked or co-crystallized with GT-B enzymes (200–202). The uracil base binds in a hydrophobic pocket formed by L197, P198, P231, and V233, which is on the second Rossmann-fold domain and facing the deep inter-domain cleft. The P231 carbonyl oxygen also hydrogen-bonds with N3 of uracil, and the G199 amide nitrogen hydrogen-bonds with the O4 carbonyl of uracil. In addition, two water molecules connect active site residues to uracil: the first water connects the G199 carbonyl oxygen and the V233 amide nitrogen to O4 of uracil, and the second water connects the G261 amide nitrogen to the O2 carbonyl of uracil. The ribose 2'- and 3'-hydroxyls both hydrogen-bond with the E281 side chain. The α- and β-phosphates hydrogen-bond with the S200 and T277 side chains, respectively. R201 also contacts both the α- and β-phosphates of UDP, albeit at a slightly longer distance (3.8 Å). R201 was shown to be critical for LpxB enzymatic activity, and this Arg side chain likely plays a catalytic role in stabilizing the negatively charged transition state of the UDP leaving group (186).

Membrane Association of LpxB

We assessed the activity of solubilized LpxB (LpxB6S) by *in vitro* enzymatic assays as well as by genetic knockout and complementation. *In vitro* activity assays

suggested that the ability of LpxB6S to catalyze the reaction of Triton X-100-solubilized substrates was completely abolished. Even after the reactions proceeded for over 17 h, the amount of UDP released was not above that of a zero enzyme control. Thus, the activity of LpxB6S was below the detection limit under these conditions (Table 3.2). In contrast, the specific activity of wild-type LpxB was measured at 6.17 ±0.53 μmol*min^{-1}*μmol^{-1} (Table 3.2). However, when the substrates were solubilized with 0.9 M 3-(1-pyridinio)-1-propanesulfonate (NDSB 201), residual lipid A disaccharide synthesis was observable for LpxB6S by TLC (Supplementary Fig. 3.3), but LpxB6S activity remained below the detection limit for UDP-release assays even with the addition of NDSB 201 (Table 3.2). These results demonstrate that LpxB6S remains catalytically competent and suggest that this solubilized form of the enzyme fails to extract lipid substrates from detergent micelles. To probe which of the residues mutated in LpxB6S are more important, we tested the *in vitro* activity of less mutated versions of LpxB. We tested two sets of triple mutants: V66S/V68S/L69S (VVL), which are located in the disordered loop (residues 62-71), and L72S/L75S/L76S (LLL), which is located in the following helix (Figure 3.2). Both triple mutants were active with detergent solubilized substrates, but they showed significantly different activities (Table 3.2): the specific activity of LpxBVVL was measured at 0.053% of wild-type while that of LpxBLLL was 0.0010% of wild-type (Table 3.2). The lower activity of the LLL mutant suggested that this later group of residues is more important for substrate binding. However, double mutants carrying any pair of L72S, L75S, and L76S retained more activity than LpxBVVL with LpxB-L72S/L75S showing the greatest decrease in activity (1.74% of wild-type versus approximately 1/3 of wild-type for the other two) (Table 3.2). These data support the

hypothesis that the hydrophobicity of this surface patch is required for extraction of lipid substrates from detergent micelles as greater decrease in overall hydrophobicity, rather than mutation of a specific residue, correlated with decrease of activity (Table 3.2).

Table 3.2: LpxB Specific Activities

LpxB	Specific Activity ±Standard Error [$\mu mol*min^{-1}*\mu mol^{-1}$]	95% Confidence Interval	Percent of Wild-type
Wild-type	$6.17 \pm 5.3*10^{-1}$	4.99 to 7.34	100
L72S/L75S	$1.08*10^{-1} \pm 2.3*10^{-2}$	$5.62*10^{-2}$ to $1.59*10^{-1}$	1.74
L72S/L76S	$1.99 \pm 3.3*10^{-1}$	1.21 to 2.76	32.2
L75S/L76S	$1.89 \pm 2.3*10^{-1}$	1.35 to 2.42	30.6
V66S/V68S/L69S (VVL)	$3.26*10^{-3} \pm 5.5*10^{-4}$	$2.02*10^{-3}$ to $4.49*10^{-3}$	$5.28*10^{-2}$
L72S/L75S/L76S (LLL)	$6.34*10^{-5} \pm 1.04*10^{-5}$	$4.03*10^{-5}$ to $8.66*10^{-5}$	$1.03*10^{-3}$
6S	Below detection limit	N/A	N/A
N316A	$2.11*10^{-2} \pm 1.5*10^{-3}$	$1.78*10^{-2}$ to $2.45*10^{-2}$	$3.42*10^{-1}$
F298E/N316A (FN)	$4.80*10^{-4} \pm 6.2*10^{-5}$	$3.33*10^{-4}$ to $6.28*10^{-4}$	$7.79*10^{-3}$
R201A	$3.16*10^{-4} \pm 2.3*10^{-5}$	$2.66*10^{-4}$ to $3.66*10^{-4}$	$5.12*10^{-3}$
FN+RA (50% FN)	$1.06*10^{-2} \pm 9*10^{-4}$	$8.50*10^{-3}$ to $1.27*10^{-2}$	$1.72*10^{-1}$
FN+RA (37% FN)	$3.63*10^{-3} \pm 3.7*10^{-4}$	$2.76*10^{-3}$ to $4.50*10^{-3}$	$5.89*10^{-2}$
FN+RA (24% FN)	$2.91*10^{-3} \pm 1.6*10^{-4}$	$2.55*10^{-3}$ to $3.27*10^{-3}$	$4.71*10^{-2}$
FN+RA (15% FN)	$2.35*10^{-3} \pm 2.6*10^{-4}$	$1.74*10^{-3}$ to $2.95*10^{-3}$	$3.81*10^{-2}$
FN+RA (7.3% FN)	$1.03*10^{-3} \pm 1.5*10^{-4}$	$6.70*10^{-4}$ to $1.38*10^{-3}$	$1.66*10^{-2}$
NA+RA1:1	$1.06*10^{-2} \pm 1.3*10^{-3}$	$7.56*10^{-3}$ to $1.35*10^{-2}$	$1.71*10^{-1}$
Wild-type (NDSB-201)	$3.42 \pm 6.0*10^{-1}$	1.994 to 4.84	55.4
6S (NDSB-201)	Below detection limit	N/A	N/A

Reactions were performed in triplicate with 1 nM to 10 μM LpxB (as appropriate per variant), 31 μM UDP-DAG, ~0.13 mM lipid X, 1 mg/mL BSA, 0.1 M Tris-HCl pH 8.0, and 0.1% Triton X-100 at ambient temperature (21°C). Reactions were quenched and UDP was quantified using a UDP-Glo Glycosyltransferase Assay kit (Promega), which quantifies free UDP by a luciferase coupled reaction. Specific activities with standard errors and 95% confidence intervals were calculated by linear regression of UDP concentrations at three or four time points for each LpxB variant in Graphpad Prism v7.03.

Activities of Previously Characterized Mutants

Wild-type LpxB and two mutants (N316A and R201A) made for this study have been previously characterized by Metzger and Raetz (186). However, the reaction conditions utilized for comparison of specific activities in this study were quite different from those utilized previously (186). Most importantly, the substrate concentrations were 31 μM UDP-DAG and ~130 μM lipid X in this study (versus 600 and 400 μM in ref.

(186)), and reactions were run at ambient temperature (~21°C) in this study (versus 30°C in ref. (186)). Thus, the specific activity of the wild-type was much lower (~1200-fold) under the conditions of this study. However, the relative activities were consistent: the specific activity of LpxB-N316A was measured at 0.34% of wild-type here and at 0.1% previously, and the specific activity of R201A was measured at 0.0051%, consistent with the <0.01% reported previously (186) (Table 3.2). The comparatively high activity of LpxB-N316A is consistent with the observed environment of N316 in the crystal structure. Rather than being positioned to interact with the substrates, the N316 side-chain hydrogen-bonds with the backbone of residues 323 and 324 and appears to stabilize the positioning of helix 15 (Figure 3.4B). LpxB-F298E/N316A (LpxBFN) showed further reduced activity (0.0078% of wild-type) (Table 3.2). Like N316, F298 is from the swapped C-terminal segment, and it makes cation-π stacking in *trans* with the side chain of critical active site residue R201 (Figures 3.3C and 3.4A). However, differential scanning fluorimetry showed no significant difference in the melting temperatures (T_m) of LpxBVVL, LpxBVVL-N316A, and LpxBVVL-R201A, and LpxBVVL-F298E/N316A had a T_m only 0.5°C lower than that of LpxBVVL (Supplementary Table 3.1). Therefore, the loss of activity in LpxB-N316A and LpxBFN is not simply related to protein unfolding.

The genetic complementation experiments generally support the conclusions of the *in vitro* assays. The genomic copy of *lpxB* was knocked out with a DNA fragment containing a kanamycin resistance gene flanked by the 3'-end of *lpxB* and the region upstream of the 5'-end in the presence of the wild-type, N316A, or F298E/N316A *lpxB* genes expressed from its own operon's promoter (203) on a plasmid.

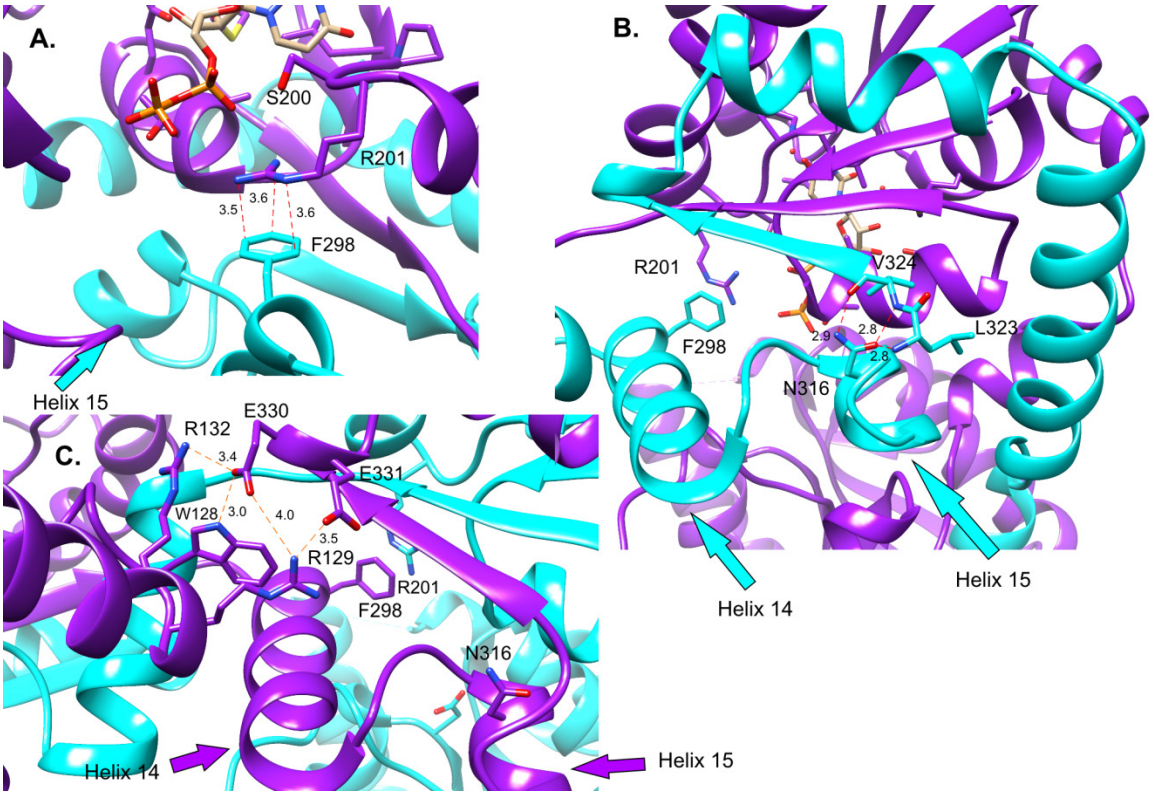


Figure 3.4: Interactions that stabilize the dimeric structure. A. Position of F298 in relation to the critical residue R201. These residues appear to be engaged in a cation-π stacking interaction with a distance of 3.6 Å between Nη1 and Cζ, a distance of 3.5 Å between Nη2 and Cε1, and a distance of 3.6 Å between Nε and Cδ2. However, the R201 side chain is not fully resolved in the electron density (Figure 3.2). The position of helix 15, which contains N316, is also visible. **B.** Position of N316 relative to the crossover helix (helix 14), showing how N316 appears to stabilize the turn from helix 15 to the final strand of the C-terminal β-sheet (strand 13). N316 hydrogen-bonds with the backbone of residues 323 and 324. The interaction distances in Ångstroms are as follows: N316 Oδ to L323 N (2.8), N316 Oδ to V324 N (2.8), and N316 Nδ to V324 O (2.9). **C.** Interactions between the N-terminal domain and the adjacent C-terminal domain. R129 forms a salt-bridge with E331 (3.5 Å between Nη1 and Oε2). E330 forms a salt-bridge with R132 (3.4 Å between Oε1 and Nη2) and a hydrogen-bond with W128 (3.0 Å between Oε1 and Nε).

In contrast, the knockout was not obtained when the gene for LpxB6S, LpxB6S-R201A, or LpxB-R201A was substituted on the plasmid. These results suggest that the hydrophobic patch (V66, V68, L69, L72, L75, and L76) is essential for productive membrane association or substrate binding. In addition, these results suggest there may be a greater difference between the activity of LpxBFN and LpxB-R201A under

physiological conditions than what was observed by UDP-release assays, wherein the specific activities were not significantly different (Table 3.2). These results appear to be consistent with TLC-based assays: LpxBFN produced a detectable amount of lipid A disaccharide after 3 h (Supplementary Fig. 3.4), but the lipid A disaccharide band in LpxB-R201A reactions only became unambiguous when the reactions were run overnight (Supplementary Fig. 3.3). Regardless, the cause of this incongruity between the genetic complementation experiments and the specific activities is unclear.

Oligomerization of LpxB in Solution

Determination of the oligomeric state of LpxB in solution has been hampered by the tendency of LpxB to form large, soluble aggregates. These aggregates are visible on a size-exclusion chromatogram (Supplementary Fig. 3.1), and this aggregated state was described as an apparent octamer that could be broken into an apparent dimer with DDM detergent by Metzger and Raetz (186). The elongated shape of LpxB further hinders the ability of size exclusion to distinguish between the monomer and dimer. In spite of this, analytical ultracentrifugation supported dimerization giving a size between 71 (right wells) and 79 kDa (left wells) (Supplementary Fig. 3.5). The total dataset did not converge, but the χ^2 residuals were minimized when the highest speed right well was excluded giving a mass of 78.9 kDa, which is most consistent with the expected 84.5 kDa dimer.

To functionally validate the C-terminal swap model, 6His-LpxB-R201A was co-expressed with LpxB-N316A or LpxBFN, which carry mutations in the C-terminus. We hypothesized that the formation of an intact active site via C-terminal swap of the R201A mutant with one of these other mutants would result in increased activity in the co-

purified protein if the second mutant were much less active than the wild-type and if the expression levels of the mutants were approximately equal. The ability of His-tagged LpxB-R201A to pull down more active LpxB-N316A in the presence of 0.2-2% Triton X-100 supports the formation of a stable oligomer (Supplementary Fig. 3.4). In addition, we mixed purified LpxB-R201A with purified LpxB-N316A or LpxBFN. However, combining 50% LpxB-R201A with 50% LpxB-N316A decreased the specific activity by approximately half indicating that the extent of heterodimer formation was not sufficient to overcome the initial ~50% dilution in activity (Table 3.2). Because these LpxB mutants lost activity when left overnight at 4^{o}C (data not shown), dimerization could not be allowed to reach equilibrium; therefore, the change in activity depends on the kinetics of subunit exchange as well as the difference in activity of mutant and intact active sites. Hence, the failure of 50% LpxB-N316A with 50% LpxB-R201A to show increased activity over 100% LpxB-N316A likely reflects slow dimer exchange since the activity of LpxB-N316A was measured well below 25% of wild-type (Table 3.2). In order to obtain a sufficiently impaired mutant to observe an increase in activity upon sub-equilibrium dimer exchange, a second mutation was made in the swapped C-terminus (F298E) producing LpxBFN as discussed above. The F298E mutation was selected based on the hypothesis that R201 stabilizes the negative charge of the UDP leaving group during catalysis (186). Because F298 already forms a cation-π interaction with R201, E298 is hypothesized to form a salt-bridge with R201 thereby balancing this critical positive charge and altering the conformation of R201 from that required for catalysis. When 50% LpxB-R201A was combined with 50% LpxBFN, the sample showed significantly more activity than freshly thawed 100% LpxBFN (Table 3.2). Increased activity was also

observed with 63%, 76%, 85%, and 92.7% LpxB-R201A, and this activity decreased with increasing LpxB-R201A concentration (Table 3.2). These data support the C-terminal swap model because the C-terminal swap would form one intact active site per R201A-F298E/N316A dimer. A R201A-F298E/N316A heterodimer with one intact active site would be expected to produce more activity than an LpxBFN homodimer because LpxBFN has only 0.0078% of wild-type activity (Table 3.2). Since minimal loss of protein stability was associated with the F298E/N316A mutations (Supplementary Table 3.1), this increase in activity is consistent with the formation of C-terminally swapped heterodimers.

3.3 Discussion

The tertiary structure of LpxB is similar to many previously characterized GT-B enzymes (96, 99, 101, 200, 202, 204); however, LpxB has a unique quaternary structure wherein the protein fold is completed by the last 87 residues of the opposite subunit in the dimer. This C-terminal swap has not been observed before in this protein superfamily, but a sialyltransferase from this superfamily involved in LPS synthesis in *Neisseria meningitidis* shows a distinct N-terminal swap of the first 130 residues (202). These unique domain arrangements likely contribute to the stability of the enzyme dimer and may potentially help coordinate activities between the two molecules.

LpxB utilizes similar active site residues as previously characterized GT-B enzymes (Table 3.3), in particular MurG (101) (18% overall sequence identity) (Supplementary Fig. 3.6). The most conserved binding motif is the use of Glu (LpxB E281) to hydrogen-bond with the 2'- and 3'-hydroxyls of the ribose of the sugar donor

substrate. In addition, S17, which was shown to be important for LpxB activity (186), is positioned similarly to the critical residue T16 in MurG (101). Finally, while the predicted catalytic base of LpxB (D98) (186) is not close in primary sequence to that of MurG (H19) (101), the side chains of these residues are close in space when the structures are overlaid (25) (Figure 3.5A).

Table 3.3: Comparison of GT-B Active Sites

Enzyme	PDB	inverting / retaining			Catalytic base			Pyro phos.		Pyro phos.	ribose
LpxB	5W8X	I	E15	S17	D98	P198	G199	S200	R201	T277	E281
MurG	1NLM	I	N/A	T16	H19	G190	G191	S192	N/A	T266	E269
Δ24PmST1	2IHZ	I	N/A	S36	D141	N/A	N/A	N/A	N/A	N/A	E338
MshA	3C4V	R	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	E324
NST	2YK7	I	N/A	N/A	D258	N/A	N/A	S322	N/A	N/A	E300
PimA	2GEJ	R	Y9	N/A	N/A	N/A	N/A	N/A	R196	N/A	E282
WaaA	2XCU	I	N/A	N/A	E31	N/A	N/A	N/A	R212	N/A	E276

Residues shown are those found to be analogous to LpxB active site residues by comparison of listed crystal structures. N/A indicates that an analogous residue could not be unambiguously assigned.

However, there are also important differences between the sugar donor binding site of LpxB and MurG. LpxB T277 appears to correspond to T266 in MurG, but T277 hydrogen-bonds with the β-phosphate while T266 hydrogen-bonds with the α-phosphate (101). Conversely, MurG S192 and LpxB S200 hydrogen-bond with the β- and α-phosphates, respectively. MurG S192 is part of a conserved GGS loop (101) while S200 is part of a conserved PGSR loop that also includes the essential residue R201 (186). Binding of the uracil moiety similarly shows variations between the enzymes: LpxB G261 is positioned similarly to MurG I245, which hydrogen-bonds uracil N3 and O4 via its backbone amide nitrogen and carbonyl (101), but the amide of LpxB G261 hydrogen bonds to uracil O2 via an intervening water. In LpxB, uracil N3 and O4 are hydrogen-bound by the backbone amides of P231, G199, and a water bound to the V233 and G199

amides. In MurG, the uracil O2 is bound by R164, a residue in the first linker between the N- and C-terminal domains (101). Finally, MurG Q288, Q289, and N128 contact the 3'- and 4'-hydroxyls of GlcNAc; however, the corresponding helix is involved in the beginning of the C-terminal swap in LpxB and is thus not positioned to contact a glucosamine bound to UDP. These structural differences suggest that LpxB also has a distinct mode of binding for the glucosamine moiety of the natural substrate, UDP-DAG. The β-hydroxymyristoyl moieties on the 2'-amine and 3'-hydroxyl of the glucosamine in UDP-DAG, likely play key roles in substrate binding. This may explain some of the differences in binding modes of the LpxB and MurG structures.

LpxB in our crystal structure appears to be in an open conformation. As observed for MurG (97, 101), PimA (99, 102), and other GT-B enzymes (98), the active site cleft of LpxB probably closes during productive substrate binding to bring the substrates into the correct conformation for nucleophilic attack by the lipid X 6'-hydroxyl on the anomeric carbon of UDP-DAG. In MurG, this involves a ~10^{o} change in the relative positions of the globular domains with the linker region acting as the hinge (101). While the intertwined dimerization mode may limit the movement of the C-terminal domain of LpxB, the contacts between the N-terminal domain and the adjacent C-terminal domain are more limited, consisting of salt bridges between R129 and E331 and between R132 and E330 and a hydrogen-bond between W128 and E330 (Figure 3.4C). Therefore, a similar hinge-like movement of the N-terminal domain may be possible. A large conformational change associated with hydrolytic removal of UMP from UDP-DAG was observed for LpxI, an enzyme that generates lipid X in a subset of Gram-negative

bacteria (82, 87). Thus, further studies are needed to capture different structural states of LpxB to fully understand its catalytic mechanism.

Previous structural studies of *E. coli* MurG revealed the presence of a surface-exposed hydrophobic patch (I75, L79, F82, W85, and W116) surrounded by a basic horseshoe (K69, K72, R80, R86, R89, and K140) (97). As mentioned above, the comparison of MurG to a homology model (160) of LpxB enabled the prediction of the LpxB hydrophobic patch. These residues form a similar hydrophobic and basic face in the two proteins; however, the exact positions of the residues is not well conserved (Figure 3.5C). PimA, which is a retaining glycosyltransferase that transfers mannose to phosphatidylinositol (16% identity to LpxB), also has a similar hydrophobic and basic face (99) (Figure 3.5C). One particularly similar feature is the flexible loop containing predicted membrane-binding hydrophobic residues followed by an amphipathic helix with a highly basic N-terminal end (Figure 3.5B). In PimA, the flexible loop includes residues 59-70, and the following surface-exposed basic residues are H76, R77, K78, K80, and K81 (99). In LpxB, the disordered loop is residues 62-71, and the basic residues are R73, R74, H77, K84, and R85 (Supplementary Fig. 3.7). Deletion of the loop or mutation of the basic residues to Ser prevented catalysis by PimA (99). Likewise, mutation of hydrophobic residues in this region of LpxB decreased the activity of LpxB. In PimA, the relevance of this helix to membrane association was further demonstrated by Förster resonance energy transfer between W82 in the helix and fluorescently labeled lipids in small unilamellar vesicles and by blue-shifting of the fluorescence emission of W82 upon association with vesicles (100).

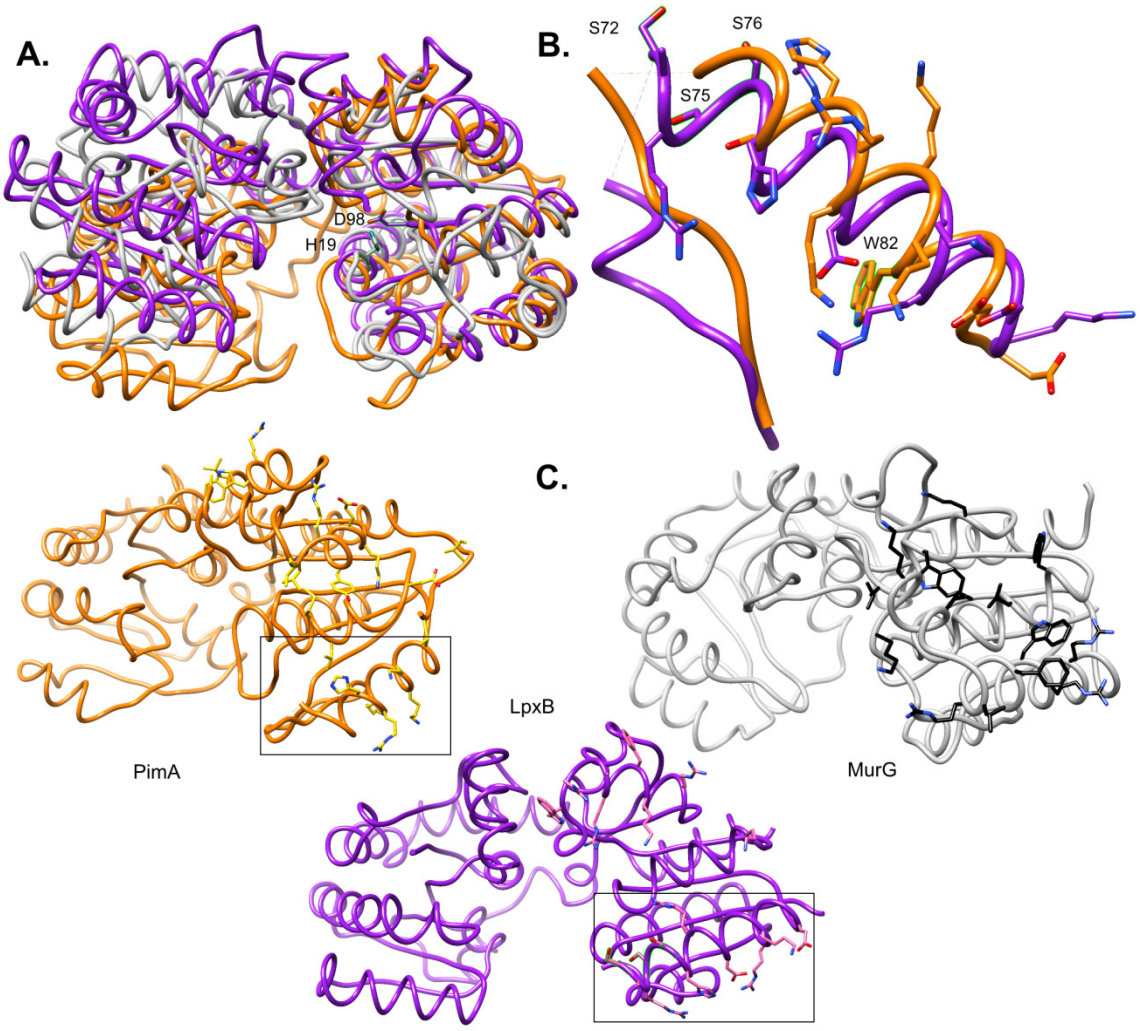


Figure 3.5: Membrane association. A. Overlay of the N-terminal domains of LpxB7S (purple) with MurG (grey) (PDB: 1F0K) (97) and PimA (orange) (PDB:2GEJ) (99). The catalytic bases of LpxB and MurG, D98 and H19 respectively, are also shown. **B.** Region of the LpxB-PimA membrane-binding face of particular interest. Deletion of the hydrophobic, flexible loop or mutation of basic residues in this helix to Ser inactivated PimA (99), and mutation of hydrophobic residues in this loop and helix to Ser (visible residues highlighted) decreased LpxB activity. In addition, W82 in the PimA helix was shown to interact with negatively charged lipid membranes (100). **C.** Side-by-side comparison of the 3 proteins, showing residues that may be important for defining the membrane association surface. While the specific residues are not well conserved, they appear to form a similar hydrophobic and basic surface in all 3 enzymes. The regions of LpxB and PimA shown in **B.** are boxed.

These data showed that the amphipathic helix associates with vesicles containing negatively charged lipids, particularly the acceptor substrate, phosphatidylinositol (100).

The membrane-association role of the amphipathic helix is likely conserved in LpxB (98). While we found LpxB6S to be inactive toward lipids in mixed micelles, the soluble enzyme remained catalytically competent when the substrates were solubilized with non-detergent sulfobetaine 201. In addition, the mutated versions of PimA and LpxB were still able to bind the nucleotide of their sugar-donor substrates (99). Therefore, the data support the hypothesis that membrane surface-active GT-B enzymes require surface-exposed hydrophobic and basic residues to extract their lipid substrates from the membrane (98, 100).

The distance (~15 Å) of the predicted catalytic base (D98) from the putative membrane-binding region (loop 62-71 and helix 4) suggests that LpxB mostly or fully extracts its lipid substrates from the membrane. The preceding enzyme, LpxH, and its alternative, LpxI, were observed to fully bind their substrates (85, 87, 154). In contrast, the following enzyme in the pathway, LpxK, was observed to primarily bind lipid head groups with the hydrocarbon tails disordered and extending into solvent (92). Regardless, LpxB, LpxH, and LpxK do not exhibit strong selectivity for chain length (80, 89). LpxB can even utilize lipid X derivatives with a single chain or a third chain attached to the hydroxyl of the 2'-β-hydroxymyristate, albeit with 2-3 orders of magnitude less activity than for lipid X (88, 205). Thus, the presence of the acyl chains may help the substrates orient productively in the active site of LpxB, but the substrates probably are not subjected to molecular ruler binding as observed in LpxA (62) and LpxD (80). Molecular docking (157) was utilized to generate plausible models for the binding of the natural donor substrate of LpxB, UDP-DAG (Figure 3.6). The models are largely consistent with the observed position of UDP in the crystal structure (Figure 3.6). In both

models shown, the lipid tails extend toward a hydrophobic groove lined by V125, W126, and W128; in addition, the anomeric carbon is exposed for nucleophilic attack from above and towards the N-terminal domain in both models. This suggests that lipid X binds second and on top of UDP-DAG, possibly with its lipid tails extending farther along the hydrophobic groove, which would explain why we were unable to obtain a plausible docking model for lipid X binding. However, the top ranked model (1 of 9) (Figure 3.6A) places the anomeric carbon 8.8 Å from the predicted catalytic base (D98). A hinge-like movement of the N-terminal domain might position D98 to accept a proton from the nucleophilic 6'-hydroxyl of lipid X. On the other hand, a lower ranked model (7 of 9) (Figure 3.6B) places the anomeric carbon 5.4 Å from D98, which should allow D98 to accept a proton from the attacking hydroxyl with minimal movement of the domain.

3.4 Conclusions and Future Directions

We have solved the structure of LpxB, revealing a new twist on the conserved GT-B fold wherein LpxB forms a dimer with the C-terminus of one subunit completing the fold of the second subunit. Soaking with UDP-GlcNAc revealed the location of the nucleotide-binding region of the active site and explained the role of previously identified conserved residues that are important for activity (186). Furthermore, our activity assays of solubilized forms of LpxB helped to provide more insight into the mode of membrane association of surface-active membrane proteins such as LpxB, PimA, and MurG. This structure should aid in the rational and computational development of new antibiotics targeting LpxB to combat the increasing problem of antibiotic resistance.

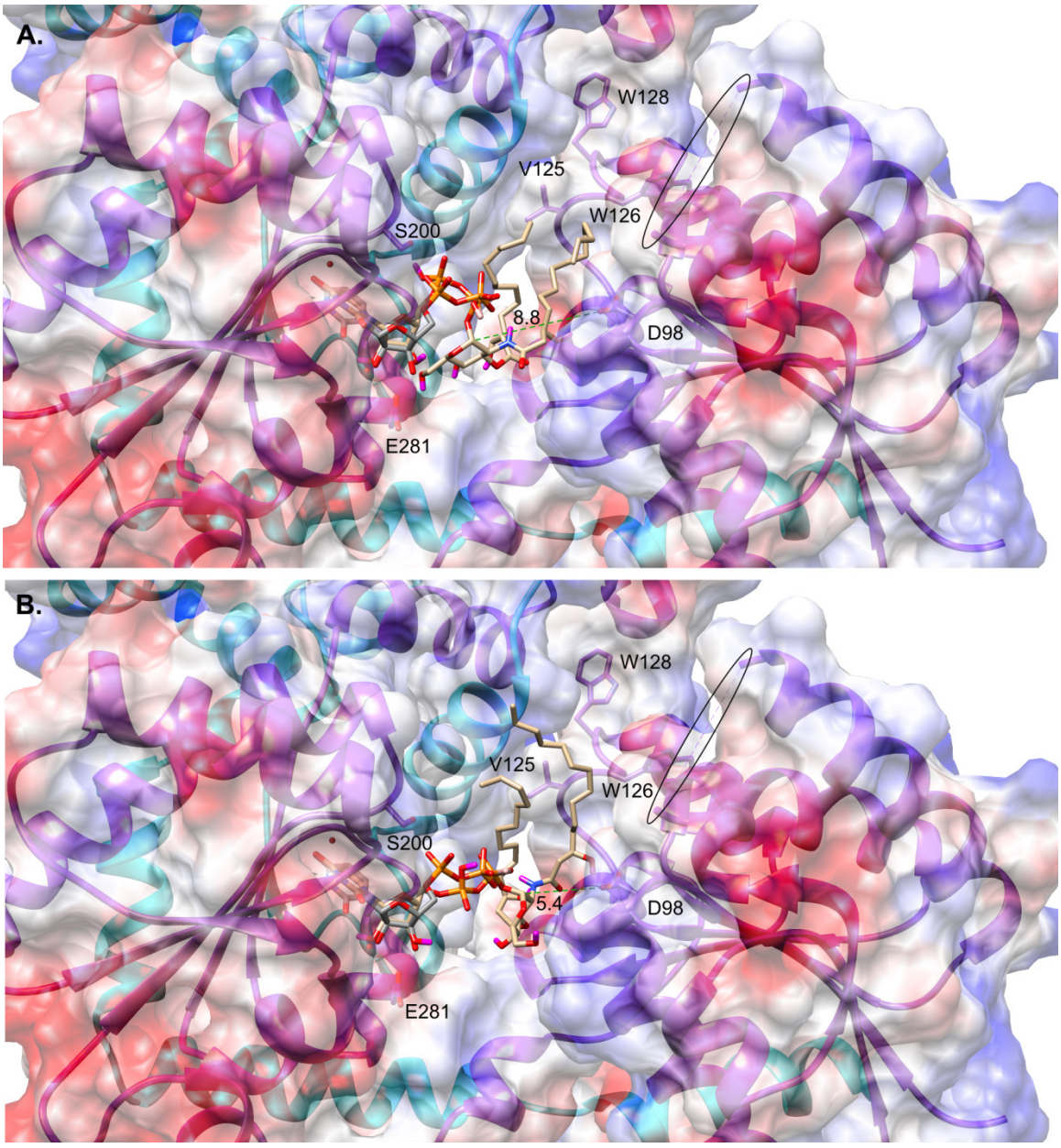


Figure 3.6: Molecular docking model of UDP-DAG binding. UDP-DAG (tan carbons) was docked into the LpxB active site cleft with AutoDock Vina (157). The software correctly identified the nucleotide-binding site overlaying reasonably well with the position of UDP (grey carbons) observed in the crystal structure. The lipid tails were predicted to extend toward a hydrophobic groove lined by V125, W126, and W128. The break in the chain corresponding to the disordered loop is circled. **A.** The top ranked model places the anomeric carbon 8.8 Å from the predicted catalytic base, D98. **B.** A lower ranked model [7 of 9] places the anomeric carbon 5.4 Å from D98, which would require less movement of the protein in order to correctly orient D98 and the 6'-hydroxyl of lipid X for proton transfer and nucleophilic attack on the anomeric carbon.

We are currently working in collaboration with Rommie Amaro's lab at University of California: San Diego to identify inhibitors of LpxB that could become antibiotic lead compounds. We will test the *in vitro* inhibition of LpxB by compounds that they identify as possible inhibitors by molecular dynamics simulations.

However, some questions on the enzymology of LpxB reamin to be answered. Most importantly, the mechanism of binding of the lipidic substrates has not been fully elucidated. Indeed, even the location of the binding site of lipid X remains vague. A crystal structure with the native substrates or lipid A dissacharide product bound would show how the substrates bind, and further activity assays could support the role of residues observed to be involved in substrate binding. In addition, the role of the putative membrane-association helix (Figure 3.5) could be tested by the mutation of basic residues within this helix (99). Finally, the role of D98 as the catalytic base (186) could be tested by further mutations of this residue. D98N would maintain the size and hydrogen-bonding capacity of D98 but prevent it from accepting a proton from the glucosamine 6'-hydroxyl while D98L would only maintain the size of D98, and both would be expected to severly decrease LpxB activity if D98 is the catalytic base. Conversely, D98H or D98C might be expected to largely maintain the activity of LpxB if D98 is the catalytic base with a possible shift in the optimal pH, particularly for D98C.

3.5 Methods

Cloning and Purification of LpxB

The LpxB coding sequence was amplified from *E. coli* BL21 cells with Phusion DNA polymerase (NEB) using the forward and reverse primers carrying BsaI (NEB) and

XbaI (NEB) restriction sites, respectively. The LpxB coding sequence was inserted into pE-SUMO expression vector (LifeSensors) to attach an N-terminally His-tagged SUMO to the N-terminus of LpxB. Following purification, the His-tagged SUMO was removed by proteolytic cleavage with the SUMO protease Ulp1. Alternatively for co-expression and purification of two mutants of LpxB, LpxB genes were inserted into pRSF-Duet-1 (Novagen) at the EcoRI (TaKaRa) and HindIII (NEB) sites of MCS1 and at the NdeI (NEB) and KpnI (NEB) sites of MCS2 thus encoding one N-terminally His-tagged protein and one untagged protein.

BL21 (DE3) cells (Lucigen) were used to express LpxB: transformed cells were grown in Miller's Luria Broth (RPI) to an OD_{600} 0.5-0.9 at 37^{o}C and then were cooled to 18^{o}C and induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) (GoldBio) overnight. For producing selenomethionine derivative proteins, the cells were grown in M9 media, and methionine synthesis was inhibited by addition of an amino acid cocktail in addition to 50-75 mg/L selenomethionine (Chem-Impex Int'l) 15 minutes prior to induction with IPTG (163). Cells were pelleted at 4,540xg (4^{o}C) for 25 min (Beckman J6-MI, JS-4.2) and then resuspended in lysis buffer (0.5 M NaCl (Fisher), 50 mM Tris-HCl (Fisher) pH 7.4, 10% glycerol (v/v) (Fisher), 5-10 mM β-mercaptoethanol (EMD Millipore)). Cells were lysed via sonication with a Branson Sonifier 450 (50% duty cycle, output 5, 6 min in three 2 min internals) while on ice. For cells expressing SUMOylated wild-type LpxB, 1% n-dodecyl-β-D-maltopyranoside (DDM) (w/v) (Anatrace D310) was added to lysate and mixed for 1 h (4^{o}C) to solubilize SUMO-LpxB. Lysate was centrifuged at 63,988xg (4^{o}C) for 45 min (Beckman Avanti J-25 I, JA-25.50) to remove cell debris. Supernatant was batch bound for 1 h (4^{o}C) to HisPur Ni-NTA resin (Thermo)

from 4-5 mL of slurry that had been equilibrated with 20 mL lysis buffer with 10 mM imidazole (Chem-Impex Int'l). Resin was collected with a gravity-flow column (room temperature) and washed with ice cold lysis buffer with 25 mM (50 mL) and 40 mM (15 mL) imidazole. SUMO-LpxB was eluted from the column with 15 mL lysis buffer with 250 mM imidazole. After addition of Ulp1, the elutant was dialyzed against 500 mL storage buffer (0.3 M NaCl, 20 mM Tris-HCl pH 7.4, 5% glycerol, 5-10 mM DTT (GoldBio), [0.05% DDM for wild-type LpxB]) overnight (4^{o}C) in 3.5 kDa cutoff tubing (Spectra/Por). Cleaved SUMO-LpxB was run back through the Ni-NTA column to remove His-tagged SUMO.

LpxB was concentrated with a 30 kDa cutoff centrifugal filter (Amicon). LpxB was then further purified via size exclusion chromatography on a Superdex 200 column (GE Healthcare) equilibrated in storage buffer. Poorly expressing wild-type LpxB was run through a 10/300 GL column. Mutant forms of LpxB with improved expression (described in *Crystallization of LpxB*) were run through a HiLoad 26/60 prep grade column. One or 2 mL fractions were collected, respectively. LpxB fractions were concentrated to 10 mg/mL, as measured by A_{280} (ε= 32805 $M^{-1}cm^{-1}$) (Nanodrop 8000 (Thermo)), for robotic crystal screening.

Protein Mass Spectrometry

The identity of purified LpxB was confirmed via liquid chromatography tandem mass spectrometry. Samples of soluble LpxB (LpxB6S described in *Crystallization of LpxB*) were submitted to Novatia (Newtown, PA). Mass/charge peaks for the expected LpxB6S molecular weight and what appears to be a cysteinylated form of LpxB6S were observed (Supplementary Fig. 3.8).

Crystallization of LpxB

Extensive sitting drop (200 nL) robotic screening (Rigaku) of wild-type LpxB revealed no successful crystallization conditions. Thus, we sought to form a crystallizable mutant of LpxB. We reasoned that the DDM micelles required to solubilize wild-type LpxB may interfere with crystal packing, so we sought to create a soluble form of LpxB. A homology model of LpxB was produced via the Phyre2 server using UDP-N-acetylglucosamine 2-epimerase from *Thermus thermophilus* HB8 (PDB: 1V4V) for sequence threading (160). Comparison of the LpxB model with the crystal structures of MurG (PDB: 1F0K and 1NLM) (97, 101), another inverting glycosyltransferase (18% identity), allowed us to identify a putative surface-exposed hydrophobic patch (V66, V68, L69, L72, L75, and L76).

Each of these residues was mutated to Ser using the Stratagene QuickChange protocol (Pfu Turbo polymerase (Agilent), DpnI (NEB)), and the expression and behavior on a size exclusion column of various combinations of these mutations was tested (Supplementary Fig. 3.1). The fully mutated form of the enzyme (LpxB6S) appeared least prone to aggregation and was selected for further robotic crystal screening. Bipyramidal and hexagonal rod shaped LpxB6S crystals grew in several conditions. The best condition (0.8 M LiCl, 32% PEG 4000, 0.1 M Tris pH 8.5) (MCSG-1 C9 Microlytic) was optimized in 2 μL hanging drops over 500 μL well solution with 1:1 8 mg/mL LpxB6S and well solution. The optimized condition for LpxB6S crystallization was 0.6-0.8 M LiCl (MP Biomedicals), 35-40% PEG 4000 (Hampton Research, HR2-529), 0.1 M Tris-HCl pH 8.6. This condition led to the formation of both crystal forms in the same drop. Unfortunately, the bipyramidal crystals that diffracted to higher resolution were

much rarer than the hexagonal rods. Additive screening (HR2-428 Hampton Research) showed that 10 mM trimethyl-ammonium chloride in the hanging drop, but not in the well solution, selected for the formation of bipyramidal crystals.

Although native LpxB6S bipyramidal crystals diffracted beyond 2.4 Å (Table 3.1), selenomethionine (SeM) derivative LpxB6S crystals that diffracted well enough for phasing by single-wavelength anomalous diffraction could not be grown. *E. coli* LpxB has 13 Met, more than adequate for phasing a 42 kDa protein. We hypothesized that one or more of these methionines was interfering with crystallization in the SeM derivative, possibly by being surface exposed and decreasing protein solubility. Thus, we further utilized the Phyre2 homology model of LpxB to predict possible culprits. Three Met residues were identified and mutated to Ser; however, only M207S preserved the stability of LpxB6S as judged by size exclusion chromatography (Supplementary Fig. 3.1). M46S and M295S increased protein aggregation. Thus, the SeM derivative of LpxB6S-M207S (LpxB7S) was selected for further crystal screening. Hexagonal rod crystals of the LpxB7S SeM derivative that diffracted to 3.43 Å (Table 3.1) were obtained in a 2 µL drop of 8 mg/mL LpxB7S (0.9 µL), well solution (0.9 µL), and 0.2 µL 100 mM trimethyl-ammonium chloride (Acros Organics) hanging over 500 µL 28% PEG 4000, 0.35 M LiBr (Sigma-Aldrich), and 0.1 M Tris-HCl pH 8.6. Switching from LiCl to LiBr was essential for obtaining harvestable LpxB7S SeM derivative crystals.

Diffraction Data Collection and Model Building

All data sets used for model building and refinement were collected at the Advanced Photon Source at Argonne National Lab on beamlines 24-ID-C and 24-ID-E at 100 K and 0.979 Å. The data were indexed, integrated, and scaled by the RAPD beamline

software (XDS (164) and CCP4 (165)) or in HKL2000 (206) (Table 3.1). The PHENIX program suite was utilized for initial phasing, model building, and refinement (166). Manual model building and further refinement were performed in Coot (167) and PHENIX Refine.

The selenomethionine anomalous diffraction data were used to calculate the initial phases and to build a crude model: AutoSol found 11 Se sites with a figure of merit of 0.348 in space group $P6_4$ and built 201 residues in 16 fragments with an R-work of 0.3786 and an R-free of 0.4692. Once the R-free of this model dropped below 0.4, this model was used as a molecular replacement model to phase the higher resolution native data in PHENIX Phaser. Sixteen fragments totaling 324 residues were built automatically in PHENIX AutoBuild in space group $P3_221$ with an R-Work of 0.2930 and an R-Free of 0.3775. The rest of the structure was built manually in Coot with iterative refinement in PHENIX Refine. In the higher resolution crystal form, there is one LpxB molecule in the asymmetric unit. Three structures were refined. The apo LpxB6S (PDB: 5W8N) structure was refined with Ramachandran statistics of 99.18% favored, 0.55% allowed, and 0.27% outliers. The apo LpxB7S structure (PDB: 5W8S) was refined with 98.63% favored, 1.10% allowed, and 0.27% outliers. The UDP-bound LpxB7S structure (PDB: 5W8X) was refined with 98.34% favored, 1.10% allowed, and 0.55% outliers.

Analytical Ultracentrifugation

Samples of LpxB6S were submitted to the Spectroscopy and Biophysics Core at University of Nebraska-Lincoln for analytical ultracentrifugation. LpxB6S (120 µL at 1.0, 0.5, and 0.25 mg/mL) was centrifuged at 5000, 7000, 9000, 12000, or 16000 rpm (2016, 3951, 6532, 11612, or 20644 xg) for 10 h in a ProteomeLab XL-A ultracentrifuge

(Beckman), and absorbance was measured at 280 nm. Only the 1 mg/mL protein produced usable data. The experiment was run in triplicate, but the middle wells contained an absorbance spike, and these wells were excluded from modeling. A solvent density of 1.0 g/mL and a protein partial specific volume of 0.73 mL/g were utilized for modeling the protein mass.

Enzymatic Assays

Genetic knockout/complementation assays of LpxB were performed in the BW25113/pKD46 strain of *E. coli* containing Lambda Red recombinase under L-arabinose induction. The cells were made chemically competent (207) and transformed with a pACYC-Duet plasmid (Novagen) containing the genes for wild-type LpxB, LpxB-R201A, LpxB-N316A, LpxB-F298E/N316A, LpxB6S, or LpxB6S-R201A. The plasmids contained the wild-type promoter for *lpxB* operon (104 bp upstream of the *fabZ* start codon) (Operon Database v3 (203)) ligated into the BamHI (NEB) and EcoRI sites and the LpxB genes ligated into the EcoRI and HindIII sites of MCS1. For experiments testing the ability of different mutants to complement each other by C-terminal swapped dimerization, a second LpxB gene was ligated into the NdeI and KpnI sites of MCS2. These strains were then grown in the presence of 1 mM L-arabinose (GeneMate), 100 mg/L ampicillin (Cayman Chemical Company), and 30 mg/L chloramphenicol (Sigma); made chemically competent (207); and transformed with a DNA fragment containing a kanamycin resistance gene flanked by 50 bp upstream of *lpxB* and the last 24 bp of *lpxB* followed by 51 more bp downstream. Chloramphenicol (30 mg/L), carbenicilin (GoldBio) (100 mg/L), and L-arabinose (1 mM) were also included in the growth step of

transformation, which was extended to 3 h. Knockout strains were selected on Miller LB agar (BD Difco) plates with 40 mg/L kanamycin (Teknova).

In vitro enzymatic assays were performed with various soluble and membrane-associated mutants of LpxB and with lipid X and UDP-DAG purified from Ni-NTA purified wild-type and D225A CcLpxI (87), respectively. For enzymatic assays, insoluble versions of LpxB were purified using Triton X-100 (Acros) instead of DDM (2% for solubilization, 0.2% for Ni-affinity column, and 0.1% for size exclusion). LpxI was expressed in C41 DE3 *E. coli* (Lucigen) and purified as for soluble LpxB except that the LpxI lysis buffer was 50 mM NaCl and 20 mM HEPES-NaOH pH 8.0. Eluted protein was then concentrated in 10 kDa cutoff centrifugal filters (Amicon). Wild-type LpxI was dialyzed against 500 mL LpxI lysis buffer, and LpxID225A was dialyzed against 50 mM NaCl and 20 mM Bis-Tris (Sigma) pH 6.0. LpxI precipitated during dialysis, but the substrates remained bound. Two volumes of methanol (Fisher) were added to the dialyzed protein, which was vortexed and then centrifuged at 3000xg (4^{o}C) for 30 min (Eppendorf 5702 R) to remove precipitate. For UDP-DAG the supernatant was concentrated in a Savant ISS110 SpeedVac (Thermo) set on medium (43^{o}C) and no further purification was performed. For lipid X, the supernatant was adjusted to pH 2 with HCl (Fisher), and dichloromethane (Sigma-Aldrich) (volume equal to methanol volume) was combined with the supernatant in a separatory funnel. Lipid X was extracted into the lower (CH_2Cl_2) phase, which was collected and then evaporated under vacuum. Lipid X residue was suspended in LpxI lysis buffer. The identities of the purified lipids were confirmed by mass spectrometry with an Agilent 1100 LC/MSD TOF mass spectrometer (G1969A) (Supplementary Fig. 3.9).

For TLC-based assays, substrates and enzyme were diluted into 2x reaction buffer (40 mM Tris-HCl pH 7.8, 0.1% Triton X-100, and 1 mg/mL BSA (Sigma)) unless otherwise noted, and 10 μL reactions were run at 30°C (Bio-Rad DNAEngine). UDP-DAG was used at 0.11 mM (as measured by UDP A_{260} (Nanodrop 8000) with ε~9.9 $mM^{-1}cm^{-1}$), and lipid X was used at ~0.13 mM as determined by lipid X depletion after reaction with 0.13 mM UDP-DAG. LpxB was used at 0.5 mg/mL, and the final pH was adjusted to 8 with 100 mM Tris-HCl pH 8.0. Reactions were quenched by spotting on HPTLC Silica gel 60 plates (EMD Millipore). Once dry (~1 h after spotting), TLC plates were run with 25/15/4/2 chloroform (Macron)/methanol/water/acetic acid (Fisher) (v/v/v/v). TLC plates were then allowed to dry at least one hour before they were sprayed with 20% sulfuric acid (Sigma-Aldrich) in ethanol (Decon) and charred with a Dual-Temp Heat Gun (Genesis) set on high (538°C). The reactants and lipid A disaccharide product appeared as darker bands. In addition, fluorescent (F_{254}) TLC Silica gel 60 plates (EMD Millipore) allowed the visualization of UDP-DAG as a fluorescent shadow by UV transmission (Bio-Rad Gel Doc EZ Imager) prior to charring.

Semi-quantitative comparisons of the relative activities of LpxB mutants were generated by performing the TLC analysis for reactions run for various times. Reactions were judged to have reached completion when the product was visible and the limiting reagent (UDP-DAG) was no longer visible. The maximum time when the reactions were not complete and the minimum time when the reactions were complete are presented as ranges (Supplementary Table 3.2). In addition, the amount of the lipids present on each plate could be compared by analyzing band intensity in ImageJ (185). TLC plate charring is an established methodology for lipid quantification though absolute quantification

requires standard curves of the lipids on interest to be run on the same plate as the samples (180, 181). The positive correlation between lipid amount and band intensity was confirmed for this study by running different amounts of UDP-DAG and lipid X and by reacting different amount of these reactants to completion with wild-type LpxB (Supplementary Fig. 3.10). The correlation was found to be positive and linear (Supplementary Fig. 3.10C and 3.10F).

While TLC-based assays have the benefit of visualizing the product of interest, they require standard curves for quantification. Thus, to produce more quantitative comparisons of LpxB activity, UDP release was quantified using a UDP-Glo Glycosyltransferase Assay kit (Promega) as described by the manufacture. A two-fold UDP (Promega) dilution series was prepared in 0.1 M Tris-HCl, 0.1% Triton X-100, and 1 mg/mL BSA. Reactions (10 μL) were run under these buffering conditions with 31 μM UDP-DAG, ~0.13 mM lipid X, and 1 nM to 10 μM LpxB at ambient temperature (~21^{o}C) in a white 384-well plate (Greiner Bio-One 781074). LpxB reactions were quenched with 10 μL UDP Detection Reagent (Promega), and the plate was shaken 30 s at 1440 rpm (1 mm amplitude) with a Spark 10M plate reader (Tecan). The luciferase coupled reaction was placed at ambient temperature for 1 hr. The resulting luminescence was then measured at 555 nm with the Spark 10M using the TR-FRET setting: 900 nm excitation, 100 μs lag time between excitation and reading, 2 ms integration of 555 nm emission, and 200 reads averaged together. Luminescence was converted to UDP concentration with concurrently run UDP standard curves prepared with 10 μL aliquots from the UDP dilution series and 10 μL UDP Detection Reagent. Reactions without LpxB served as a negative control: UDP detected at or below the concentration in these

samples was assumed to result from UDP-DAG hydrolysis rather than formation of lipid A disaccharide. Of note, the UDP concentration of negative controls only measured significantly above 0 in overnight reactions, and never measured above 0.5 μM. Reactions and standard curve samples were prepared in triplicate. Reactions were quenched at various time points to obtain at least 3 time points per LpxB variant within the early, linear product accumulation range of the reactions. Specific activities with standard errors and 95% confidence intervals were obtained by linear regression analysis of these data in Graphpad Prism v7.03.

Differential Scanning Fluorimetry

DSF was performed with SYPRO Orange (Life Technologies) and soluble versions of LpxB diluted into soluble LpxB storage buffer to final concentrations of 80x SYPRO Orange (1.6% DMSO) and 0.8 mg/mL LpxB. Aliquots (40 μL) were heated from 20 to 95^{o}C on a C1000 Touch Thermal Cycler and fluorescence was measured with the FRET channel of a CFX96 Real-Time System (Bio-Rad). The step size of the temperature ramp was 0.5^{o}C, and the time at each temperature before fluorescence scanning was 30 s. Melting curves were analyzed in Bio-Rad's CFX Manager software to get the melting temperatures (T_m) of the LpxB mutants.

Molecular Docking

UDP-DAG was docked into the LpxB active site cleft with AutoDock Vina 1.1.2 (157). The apo LpxB7S structure (PDB: 5W8S) and the UDP-DAG molecule were prepared for docking in AutoDockTools (182, 183) from the MGLTools 1.5.6 suite. For docking, the active site was defined as a box centered at coordinates (-36.232, 20.151,

12.865) with x, y, z lengths of 24, 36, and 32 Å. Exhaustiveness was set to 8. Nine models were output.

Data Availability

The coordinates and structure factors of the crystals structures generated from this research are available at the Protein Data Bank under accession numbers 5W8N (LpxB6S), 5W8S (LpxB7S), and 5W8X (LpxB7S bound to UDP). All other relevant data are available in this article and its Supplementary Information files, or from the corresponding author upon request.

3.6 Supplementary Information

Supplementary Table 3.1: LpxB Melting Temperature

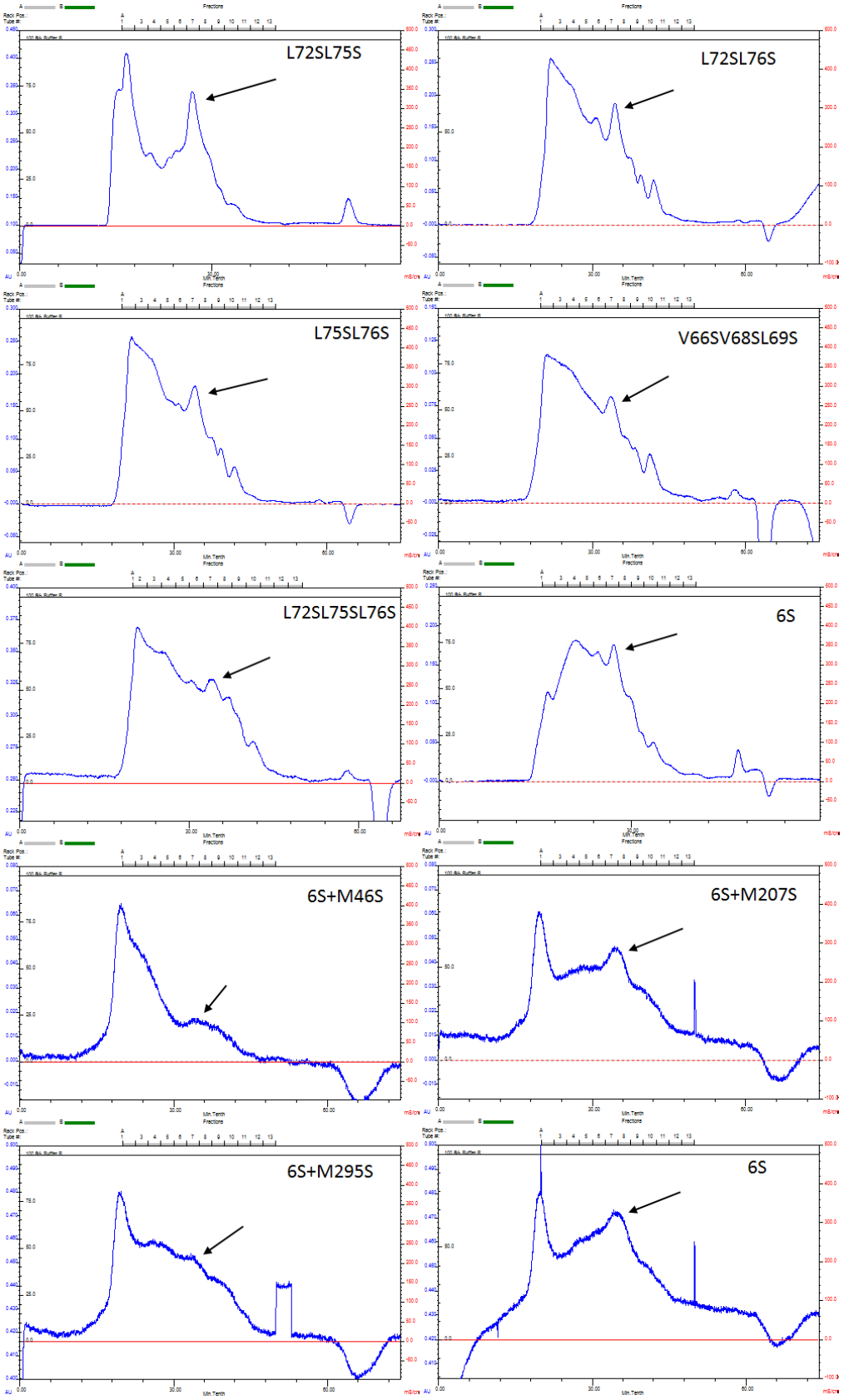
LpxB	T_m (°C)
V66S/V68S/L69S (VVL)	49.0 ± 0.4
VVL-R201A	49.5 ± 0.4
VVL-N316A	49.2 ± 0.5
VVL-F298E/N316A	48.5 ± 0.0

Melting was observed by SYPRO Orange fluorescence. T_m values are the temperatures at the inflection points of the melting curves plus or minus the standard deviations (n= 4).

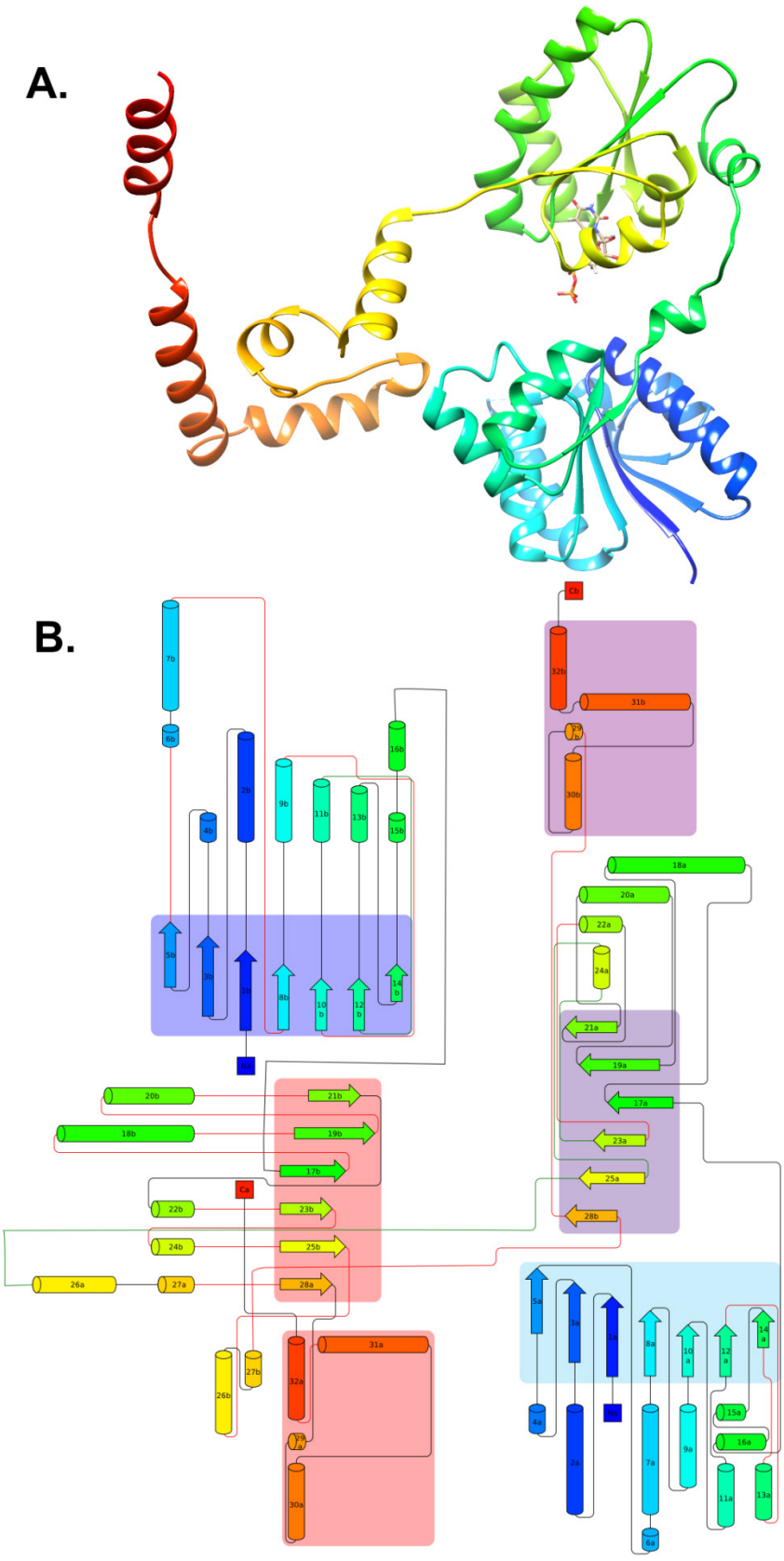
Supplementary Table 3.2: Semi-quantitative Comparison of LpxB Activities by TLC

LpxB	Time to Reaction Completion
wild-type	1 min<t<5 min
L72S/L75S	10 min<t<30 min
L72S/L76S	1 min<t<5 min
L75S/L76S	1 min<t<5 min
V66S/V68S/L69S (VVL)	1 h<t<2 h
L72S/L75S/L76S (LLL)	5 h<t< 6 h
6S	18.75 h*<t
N316A	30 min<t<1 h
F298E/N316A	15.5 h‡<t
R201A	18.75 h<t
6S-R201A	No activity detected

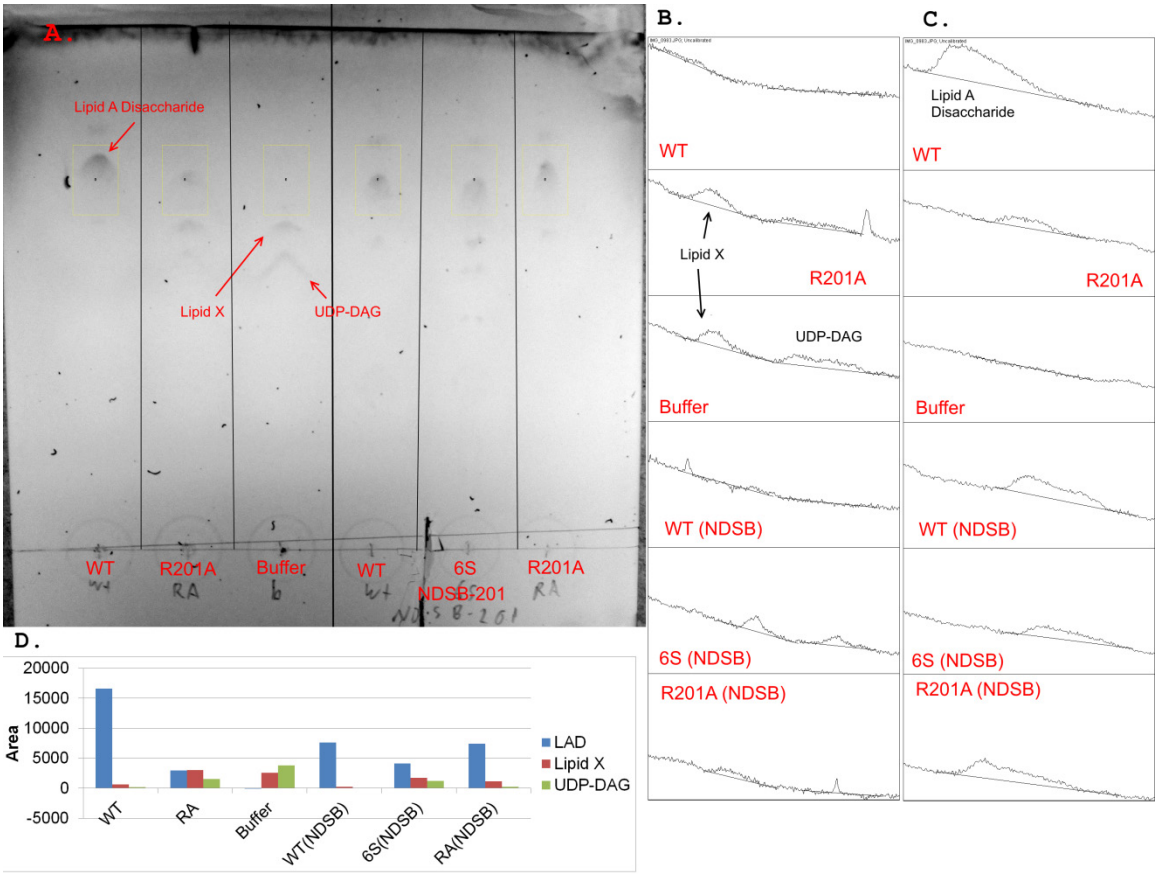
Reactions were performed with 0.5 mg/mL LpxB, 0.11 mM UDP-DAG, ~0.13 mM lipid X, 0.5 mg/mL BSA, 0.05-0.06% Triton X-100 and 0.1 M Tris-HCl pH 8 at 30°C. *LpxB6S showed no activity under standard conditions but was active when 0.9 M NDSB 201 was added. ‡Reactions with LpxBFN appeared mostly complete by 4.5 h, but were never observed to go to completion.


L72SL75S
L72SL76S
L75SL76S
V66SV68SL69S
L72SL75SL76S
6S
6S+M46S
6S+M207S
6S+M295S
6S

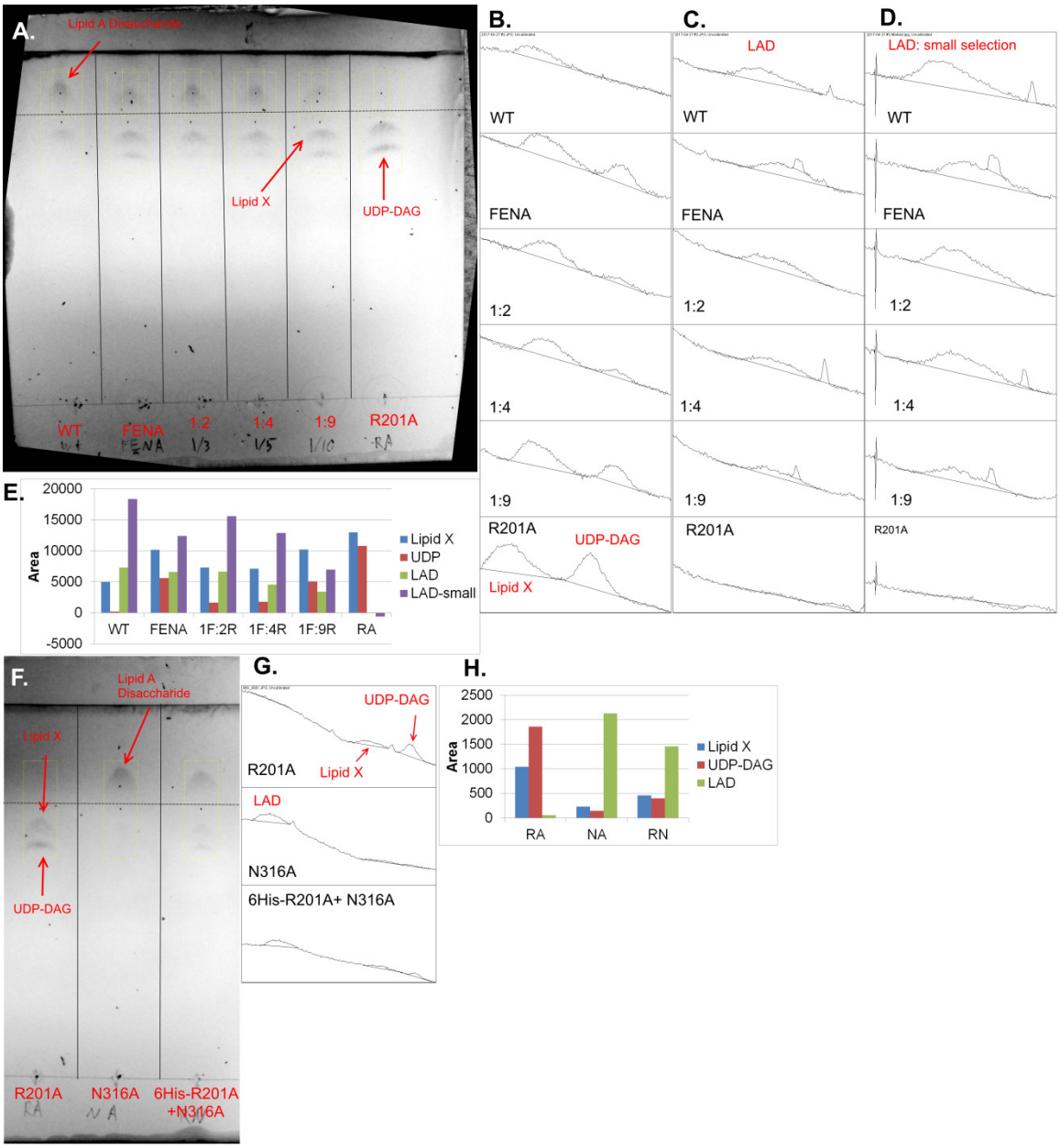
Supplementary Fig. 3.1: Size exclusion chromatography. Ni-NTA purified LpxB mutants were run on 10/300 GL Superdex 200 columns (GE Healthcare). The top 6 UV traces are from protein run on an "Increase" model column at 0.5 mL/min. The bottom 4 UV traces are from protein run on the standard model column. Crystallizable LpxB appeared as a peak in fractions 7-8 (indicated by arrows). Large amounts of LpxB ended up in the void/soluble aggregate fractions preceding fraction 7. The quality of various mutants was judged by the peak heights of the aggregated and fraction 7-8 peaks. LpxB6S has relatively little protein in the void peak compared to soluble forms with fewer mutations. M207S preserves the behavior of LpxB6S while M46S and M295S lead to more protein aggregation.


A.
B.

Supplementary Fig. 3.2: LpxB forms an intertwined dimer. A. shows the isolated UDP-bound LpxB monomer. Spectrum coloring ends with red at the C-terminus. **B.** shows the 2D depiction of the LpxB dimer generated with Pro-origami (199). The rainbow coloring of secondary structural elements is similar to that in **A**, and the numbering of secondary structural elements is sequential for each subunit (a and b). Colored boxes indicate strands (arrows) in the same sheet and the final helices (cylinders) that wrap around the bottom of the other subunit.

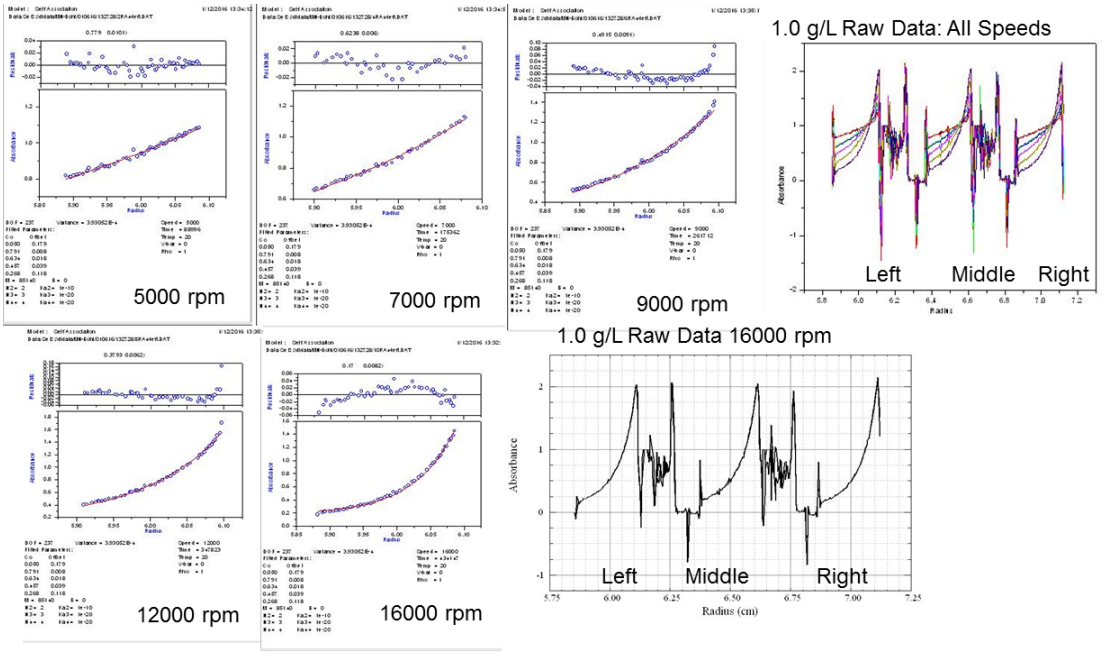


Supplementary Fig. 3.3: Activity of LpxB6S mutant. A. Products from overnight (18.75 h) reactions under standard conditions (left 3) or with 0.9 M NDSB 201 (right 3) were run on a Silica gel 60 HPTLC plate. Charring shows product in the all lanes except the no enzyme control. Lipid X is visible in all lanes, and UDP-DAG is visible in the left LpxB-R201A lane, the LpxB6S lane, and the negative control lane. These results were replicated once. The intensity of the bands in **A** were analyzed in ImageJ (185): defined lanes were plotted as distance versus intensity, and the area under the peak was determined. **B.** shows the curves for the reactant bands, and **C.** shows the curves for the product band. **D.** shows a graphical depiction of the areas under the peaks in **B** and **C**.

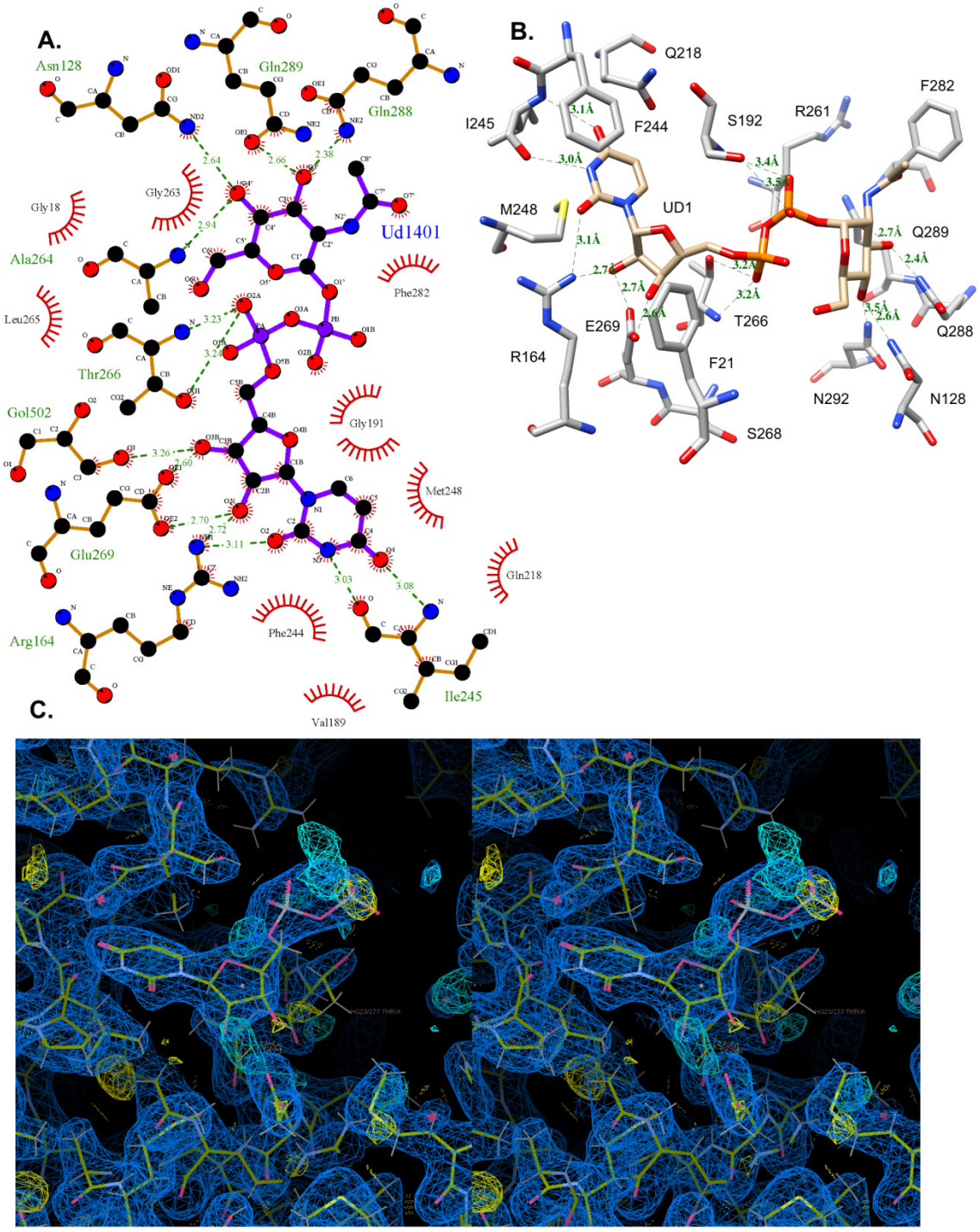


Supplementary Fig. 3.4: Dimerization of R201A with F298E/N316A and N316A. A. Three hour reactions included wild-type LpxB, LpxB-R201A, and LpxBFN combined in shown ratios with LpxB-R201A (constant total protein). Consistent with the UDP-release assays, UDP-DAG bands are fainter and product bands appear darker in 1:2 and 1:4 lanes than in the 100% LpxBFN lane. The increase in activity of the 1:2 mixture relative to 100% LpxBFN was replicated 4 times. TLC analysis of the 1:4 and 1:9 mixtures was not repeated. **B.** Reactant band intensity was analyzed in ImageJ (185): lanes were plotted as distance versus intensity and the areas under the peaks were determined. **C.** The same analysis of product bands. The irregular shape of these bands gives poor peak resolution, making quantification more difficult. **D.** When only the central, darkest potion of the product bands is selected, peak resolution is improved. **E.** The

quantification of the peak areas. The decrease in lipid X and UDP-DAG intensities 1:2 and 1:4 mixtures support the increase in activity produced by combining LpxBFN and LpxB-R201A observed in UDP-release assays. As in the UDP-release assays, activity appears to decrease with increasing percent LpxB-R201A. When the product band intensity is quantified with the smaller, central selection, the wild-type and 1:2 mixture lanes show a corresponding increase in product band intensity. However, when the full band was selected, there was little difference between wild-type, LpxBFN, and the 1:2 mixture, and the 1:4 mixture even showed decreased product band intensity. The 1:9 mixture shows weaker product band intensity in either case. Regardless, these data support the conclusion from the UDP-release assays that LpxBFN and LpxB-R201A form a complementary dimer. **F.** One hour reactions of LpxB-R201A, LpxB-N316A, and co-purified His-tagged LpxB-R201A with untagged LpxB-N316A. While LpxB-R201A did not produce a detectable amount of product, LpxB-N316A and the co-purified proteins show production of lipid A disaccharide with corresponding depletion of reactants indicating that LpxB-R201A can pull down LpxB-N316A. These results were replicated five times. **G. H.** Analysis was performed as for the plate in **A.**

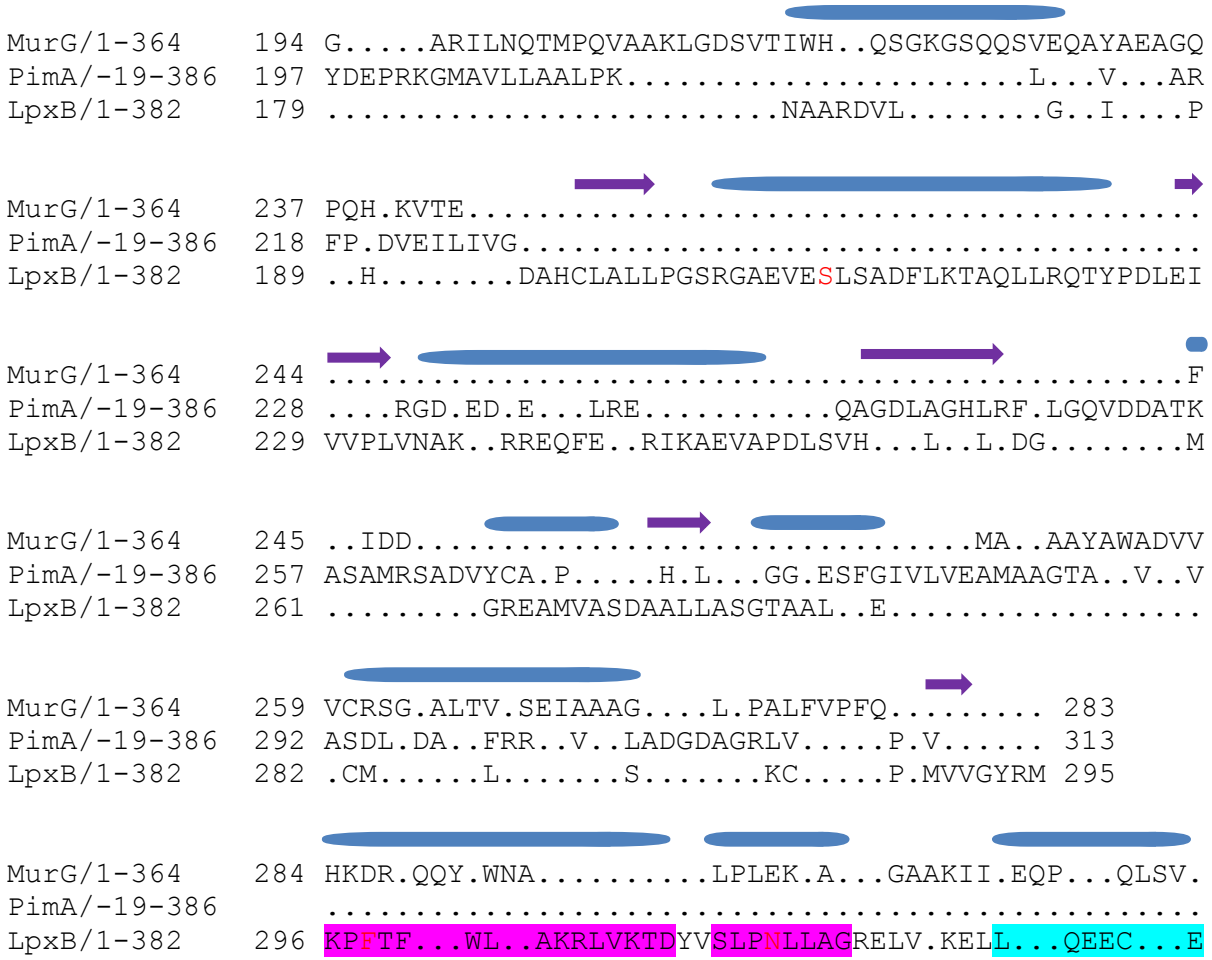


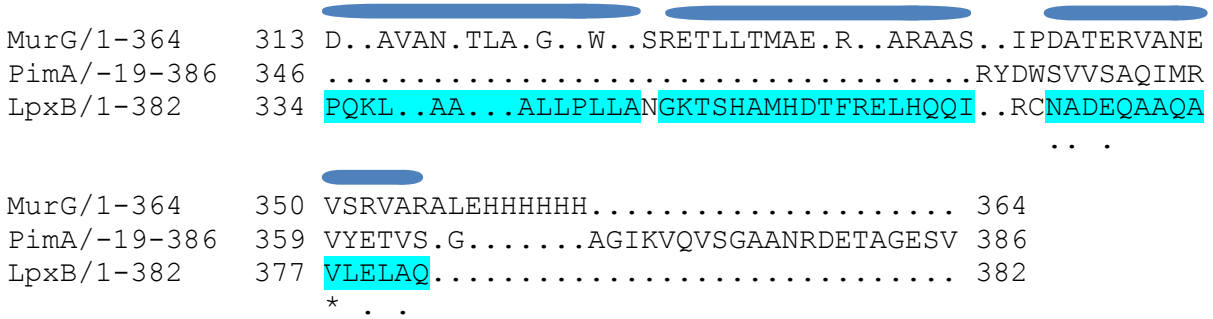
Supplementary Fig. 3.5: Analytical ultracentrifugation. Raw and fitted analytical ultracentrifugation data are shown. LpxB6S (120 µL) was centrifuged at 5000, 7000, 9000, 12000, or 16000 rpm for 10 h in a ProteomeLab XL-A ultracentrifuge. Absorbance was measured at 280 nm. Fitting of distance versus absorbance data is shown for the left wells. Fitting with a solvent density of 1.0 g/mL and a protein partial volume of 0.73 mL/g gave a protein mass of 71-79 kDa, most consistent with the expected 84.5 kDa dimer.



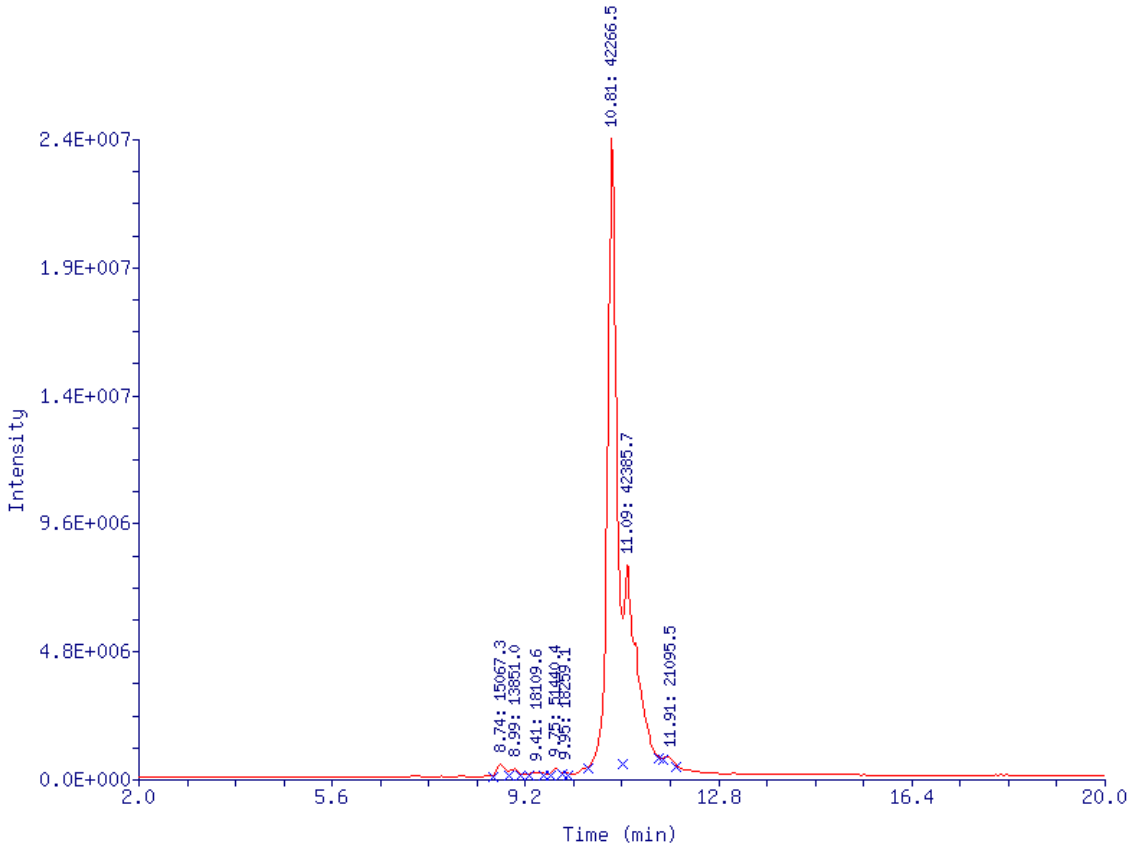
Supplementary Fig. 3.6: MurG active site. A. B. The MurG active site from PDB entry 1NLM (101) is provided for comparison to Figure 3. **C.** Stereo view of LpxB UDP-binding pocket. The 2mFo-DFc map (blue mesh) is shown at 0.4317 e/Å^3 (0.99 RMSD), and the mFo-DFc map (cyan mesh for positive and yellow mesh for negative) is shown at 0.45 e/Å^3 (3.01 RMSD).


MurG/1-364 1 MMSGQG.....KRLMVMA.G.....GTGG
PimA/-19-386 -19 .....MGSSHHHHHHSSGLVPRGSHMRIGMVCPYSFDVPG..G
LpxB/1-382 1 ......MTEQRP.....LTIALVAGE.....T..S
MurG/1-364 19 HVFPGLAVAHHLM.A..Q.GWQVRWLGTADRMEA.DLVPKHGI.EIDFI.
PimA/-19-386 17 VQSHVLQLAEVLRDA....GHEVSVLAPA...SPHVK.L..PD.YVVSG.
LpxB/1-382 18 GDILGAGLIRAL..KEHVPNARFVGVA.....G..PRMQ.AEGCEAWYEM
MurG/1-364 62 R.IS.....GLRGKGIKALIAAPLRIFNAWR
PimA/-19-386 55 GK.A.....VPIPYNGSVARLRF.....G..PATHR
LpxB/1-382 58 EE.LAVMGISESSGR.....S.....RRSSHIRA
MurG/1-364 87 QARAIMKAYKPDVVLGMGG...YVSGPGGLAAWSLGIPVVLHEQNGI...
PimA/-19-386 78 KVKKWIAEGDFDVLHIHEPNAPSLSMLALQAAE...GPIVATFHT.STTK
LpxB/1-382 81 DLTKRFGELKPDVFVGIDAP.DF.NITLEGNLKKQGIKTIHYVS..P...
MurG/1-364 131 .AGLTN..K....W.LAK..IATKVMQAFPGA..FPN.....AEVVGNP
PimA/-19-386 124 SLT...LSVFQGILRPYHE.KIIGRIAVSDLARRWQMEALGSDAVEIPNG
LpxB/1-382 124 .SVW....AWRQKRVFKIGRATDLVLAFLPFEKAFYDK.YNVPCRFIGHT
MurG/1-364 163 ..VRTD...VLALPLPQQRLAGREGPVRVL.....VVGGSQ
PimA/-19-386 170 ..VDVASF.A.....DAPLLDGYPREGRTVLFLGR
LpxB/1-382 168 MADA....MP.....L...D.....PDK.....

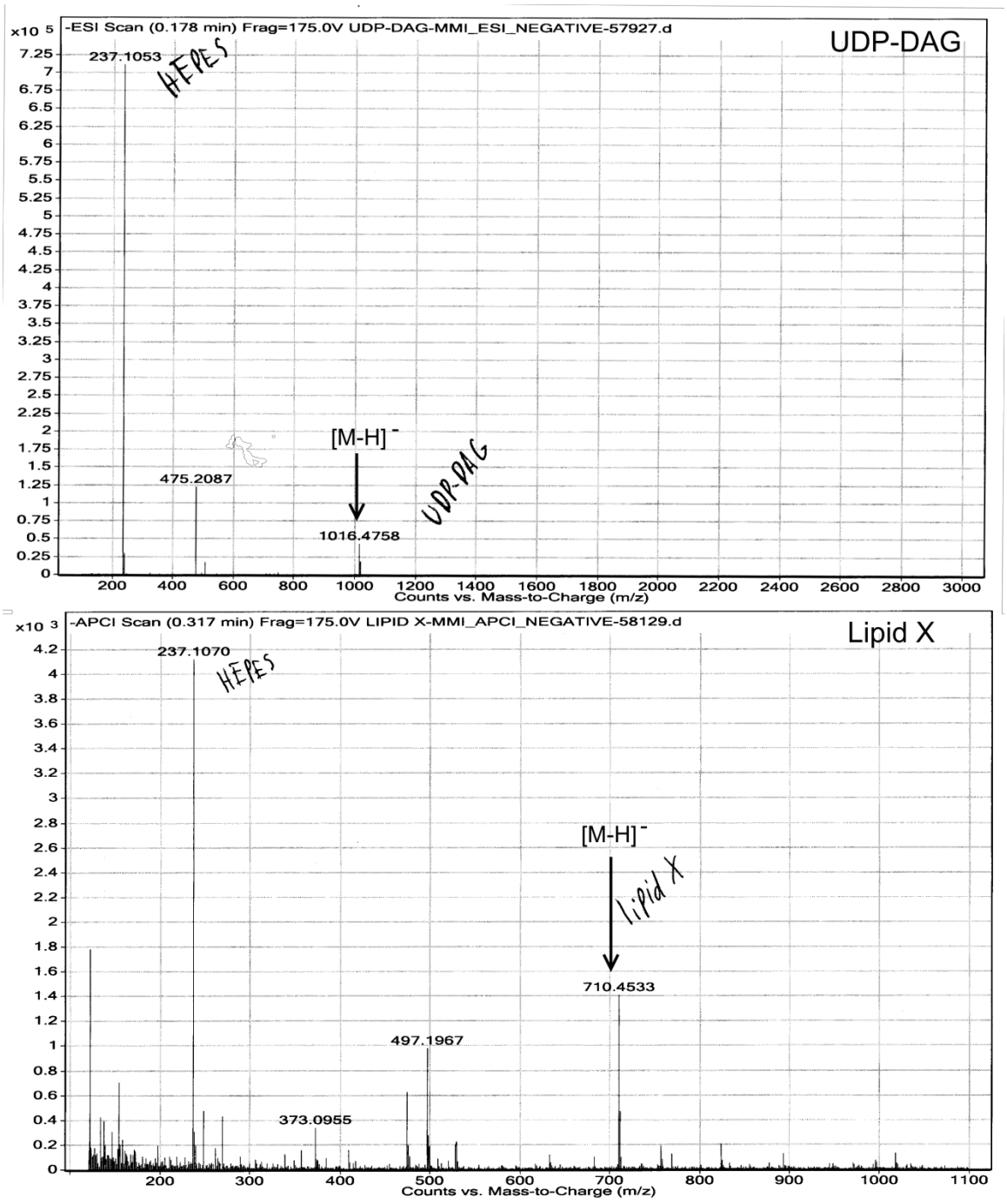

MurG/1-364 194 G.....ARILNQTMPQVAAKLGDSVTIWH..QSGKGSQQSVEQAYAEAGQ
PimA/-19-386 197 YDEPRKGMAVLLAALPK.....L...V...AR
LpxB/1-382 179 .....NAARDVL.....G..I....P
MurG/1-364 237 PQH.KVTE.....
PimA/-19-386 218 FP.DVEILIVG.....
LpxB/1-382 189 ..H.....DAHCLALLPGSRGAEVESLSADFLKTAQLLRQTYPDLEI
MurG/1-364 244 .....F
PimA/-19-386 228 ....RGD.ED.E...LRE.....QAGDLAGHLRF.LGQVDDATK
LpxB/1-382 229 VVPLVNAK..RREQFE..RIKAEVAPDLSVH...L..L.DG.....M
MurG/1-364 245 ..IDD.....MA..AAYAWADVV
PimA/-19-386 257 ASAMRSADVYCA.P.....H.L...GG.ESFGIVLVEAMAAGTA..V..V
LpxB/1-382 261 .....GREAMVASDAALLASGTAAL..E.....
MurG/1-364 259 VCRSG.ALTV.SEIAAAG....L.PALFVPFQ..... 283
PimA/-19-386 292 ASDL.DA..FRR..V..LADGDAGRLV.....P.V..... 313
LpxB/1-382 282 .CM.....L.....S.....KC.....P.MVVGYRM 295
MurG/1-364 284 HKDR.QQY.WNA.....LPLEK.A...GAAKII.EQP...QLSV.
PimA/-19-386 .....
LpxB/1-382 296 KPFTF...WL..AKRLVKTDYVSLPNLLAGRELV.KELL...QEEC...E



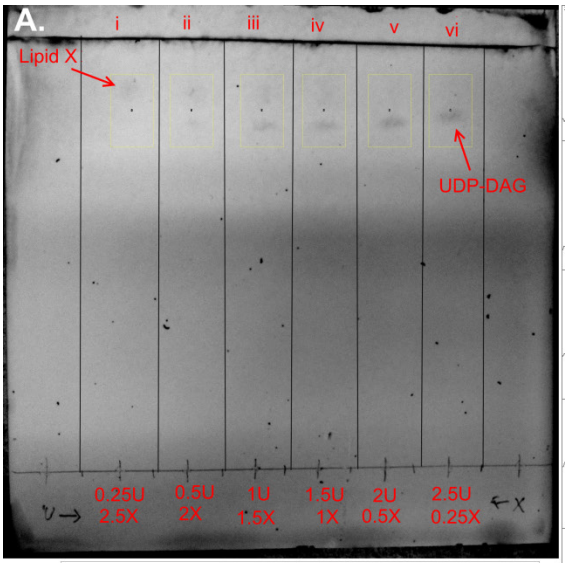
Supplementary Fig. 3.7: Structural alignment of LpxB (PDB: 5W8S), MurG (PDB: 1F0K), and PimA (PDB: 2GEJ). Purple arrows and blue ovals show the positions of β-sheets and α-helices in the LpxB sequence. Yellow highlighting marks the disordered loops in LpxB and PimA, and green highlighting marks the amphipathic helices involved in membrane association (99, 100). Magenta highlighting marks helices 14 and 15 in LpxB which contain F298 and N316, respectively. Cyan highlighting marks the final 3 helices in LpxB. Red lettering marks the locations of point mutations made in LpxB.

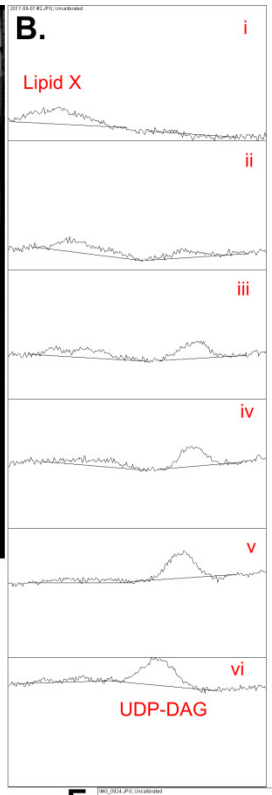


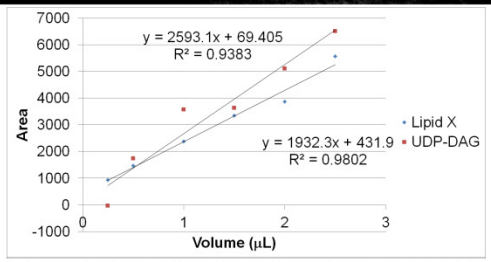
Supplementary Fig. 3.8: Liquid chromatography-mass spectrometry. LCMS of LpxB6S gives a peak corresponding to a mass of 42266.5 Da, which corresponds well with the calculated molecular mass of 42270.01 Da from the ProtParam server <http://web.expasy.org/protparam/>. A second peak at 42385.7 (119.2 Da larger) probably corresponds to a cyteinylated form of LpxB. Indeed, electron density at C367 of the unbound LpxB7S structure suggests some modification (data not shown). This is likely a result of oxidation stress in the *E. coli.*

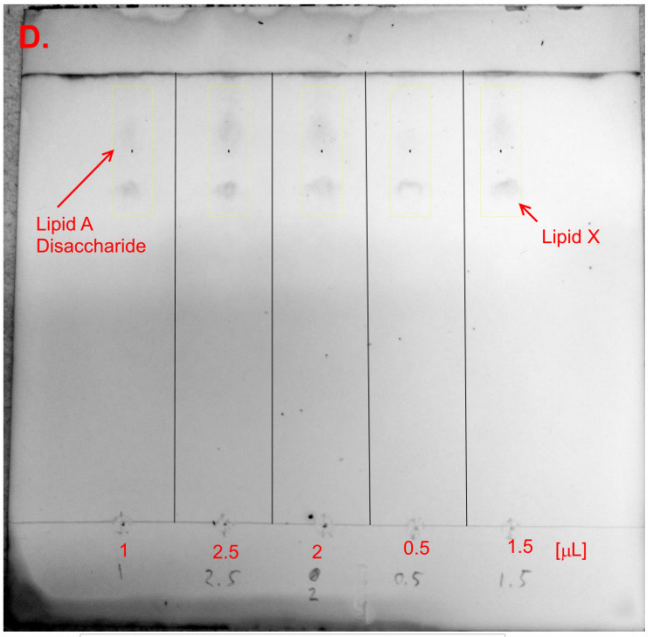


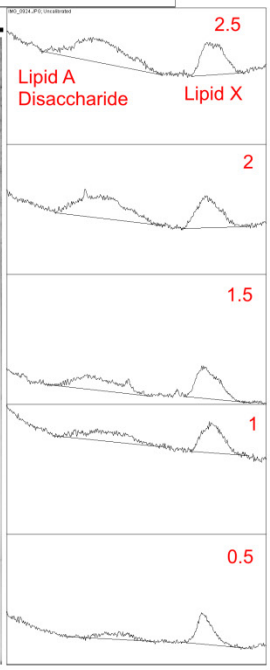
Supplementary Fig. 3.9: Mass spectra of LpxB reactants. UDP-DAG and Lipid X were extracted from CcLpxI, wild-type and D225A respectively (87). The identity of the purified lipids was confirmed by their $[M-H]^-$ peaks, which correspond well with the predicted exact masses of lipid X (711.43 Da) and UDP-DAG (1017.46 Da). The mass spectra were collected in negative ion mode with direct infusion of the diluted lipids and ionization by electrospray or atmospheric pressure chemical ionization.

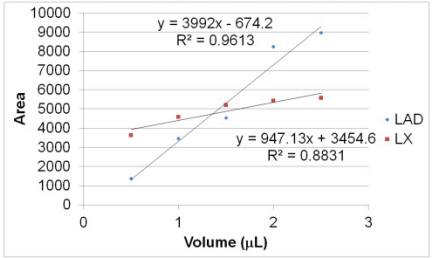

A.
i
ii
iii
iv
v
vi
Lipid X
UDP-DAG
0.25U 2.5X
0.5U 2X
1U 1.5X
1.5U 1X
2U 0.5X
2.5U 0.25X


B.
i
Lipid X
ii
iii
iv
v
vi
UDP-DAG


C.
y = 2593.1x + 69.405
R² = 0.9383
y = 1932.3x + 431.9
R² = 0.9802
Area
Volume (µL)
Lipid X
UDP-DAG


D.
Lipid A Disaccharide
Lipid X
1
2.5
2
0.5
1.5
[µL]


E.
2.5
Lipid A Disaccharide
Lipid X
2
1.5
1
0.5


F.
y = 3992x - 674.2
R² = 0.9613
y = 947.13x + 3454.6
R² = 0.8831
Area
Volume (µL)
LAD
LX

Supplementary Fig. 3.10: Linear relationship between lipid amount and band intensity. A. Different volumes [μL] of UDP-DAG (0.31 mM) and lipid X (~0.31 mM) were run on an HPTLC Silica gel 60 plate. Once dry, the plate was charred with 20% sulfuric acid in ethanol and a heat gun. This plate was replicated once. **B.** Band intensities were analyzed in ImageJ (185): lanes were plotted as distance versus intensity and the areas under the peaks were determined. **C.** Volumes run were plotted versus peak areas, and the data were fitted by a line in MS Excel. Both reactants show increasing band intensity with increasing lipid amount, and the relationships are quite linear: R^2 of 0.9802 for lipid X and 0.9383 for UDP-DAG. **D.** Different volumes [μL] of UDP-DAG (0.31 mM) and lipid X (1.3 mM) were reacted to completion with 0.5 g/L wild-type LpxB in 0.1 M Tris-HCl pH 8.0, 0.1% Triton X-100, and 1 mg/mL BSA at 30^{o}C for 1 hr. TLC of products was performed as in **A.** This plate was not repeated. **E.** Band intensities of lipid A disaccharide product and remaining lipid X were analyzed as in **C. F.** Volumes reacted were plotted versus the band intensities, and the data were fitted by a line in MS Excel. Band intensities increase with increasing lipid amounts. The lipid A disaccharide data are quite linear (R^2 of 0.9613), and the lipid X data are reasonably linear (R^2 of 0.8831).

BIBLIOGRAPHY

1. Raetz, C. R. H., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700
2. Malinverni, J. C., and Silhavy, T. J. (2009) An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8009–8014
3. Maldonado, R. F., Sá-Correia, I., and Valvano, M. A. (2016) Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* **40**, 480–493
4. Bojkovic, J., Richie, D. L., Six, D. A., Rath, C. M., Sawyer, W. S., Hu, Q., and Dean, C. R. (2015) Characterization of an Acinetobacter baumannii lptD deletion strain: permeability defects and response to inhibition of lipopolysaccharide and fatty acid biosynthesis. *J. Bacteriol.* **198**, 731–741
5. Reynolds, C. M., and Raetz, C. R. H. (2009) Replacement of lipopolysaccharide with free lipid A molecules in Escherichia coli mutants lacking all core sugars. *Biochemistry (Mosc.)*. **48**, 9627–9640
6. Beceiro, A., Moreno, A., Fernández, N., Vallejo, J. A., Aranda, J., Adler, B., Harper, M., Boyce, J. D., and Bou, G. (2014) Biological cost of different mechanisms of colistin resistance and their impact on virulence in Acinetobacter baumannii. *Antimicrob. Agents Chemother.* **58**, 518–526
7. Opiyo, S. O., Pardy, R. L., Moriyama, H., and Moriyama, E. N. (2010) Evolution of the Kdo2-lipid A biosynthesis in bacteria. *BMC Evol. Biol.* 10.1186/1471-2148-10-362
8. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) Lipid A modification systems in gram-negative bacteria. *Annu. Rev. Biochem.* **76**, 295–329
9. Miyake, K. (2006) Roles for accessory molecules in microbial recognition by Toll-like receptors. *J. Endotoxin Res.* **12**, 195–204
10. Kim, J.-I., Lee, C. J., Jin, M. S., Lee, C.-H., Paik, S.-G., Lee, H., and Lee, J.-O. (2005) Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *J. Biol. Chem.* **280**, 11347–11351
11. Klein, G., Lindner, B., Brade, H., and Raina, S. (2011) Molecular basis of lipopolysaccharide heterogeneity in Escherichia coli: envelope stress-responsive regulators control the incorporation of glycoforms with a third 3-deoxy-α-D-manno-oct-2-ulosonic acid and rhamnose. *J. Biol. Chem.* **286**, 42787–42807
12. Klein, G., Müller-Loennies, S., Lindner, B., Kobylak, N., Brade, H., and Raina, S. (2013) Molecular and structural basis of inner core lipopolysaccharide alterations in Escherichia coli: incorporation of glucuronic acid and phosphoethanolamine in the heptose region. *J. Biol. Chem.* **288**, 8111–8127
13. Ohto, U., Fukase, K., Miyake, K., and Satow, Y. (2007) Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science*. **316**, 1632–1634
14. Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., and Lee, J.-O. (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature*. **458**, 1191–1195

15. Shimazu, R., Akashi, S., Ogata, H., Nagai, Y., Fukudome, K., Miyake, K., and Kimoto, M. (1999) MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4. *J. Exp. Med.* **189**, 1777–1782
16. Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Van Huffel, C., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., Freudenberg, M., Ricciardi-Castagnoli, P., Layton, B., and Beutler, B. (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science*. **282**, 2085–2088
17. Hall, M. J., Williams, S. N., DeFrances, C. J., and Golosinskiy, A. (2011) Inpatient care for septicemia or sepsis: a challenge for patients and hospitals. *NCHS Data Brief*. **62**, 1–8
18. John, C. M., Phillips, N. J., Stein, D. C., and Jarvis, G. A. (2017) Innate immune response to lipooligosaccharide: pivotal regulator of the pathobiology of invasive Neisseria meningitidis infections. *Pathog. Dis.* 10.1093/femspd/ftx030
19. Ventola, C. L. (2015) The antibiotic resistance crisis: part 2: management strategies and new agents. *P T Peer-Rev. J. Formul. Manag.* **40**, 344–352
20. Loppnow, H., Brade, H., Dürrbaum, I., Dinarello, C. A., Kusumoto, S., Rietschel, E. T., and Flad, H. D. (1989) IL-1 induction-capacity of defined lipopolysaccharide partial structures. *J. Immunol. Baltim. Md 1950*. **142**, 3229–3238
21. Ulmer, A. J., Heine, H., Feist, W., Kusumoto, S., Kusama, T., Brade, H., Schade, U., Rietschel, E. T., and Flad, H. D. (1992) Biological activity of synthetic phosphonooxyethyl analogs of lipid A and lipid A partial structures. *Infect. Immun.* **60**, 3309–3314
22. Rietschel, E. T., Kirikae, T., Schade, F. U., Mamat, U., Schmidt, G., Loppnow, H., Ulmer, A. J., Zähringer, U., Seydel, U., and Di Padova, F. (1994) Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **8**, 217–225
23. Golenbock, D. T., Hampton, R. Y., Qureshi, N., Takayama, K., and Raetz, C. R. (1991) Lipid A-like molecules that antagonize the effects of endotoxins on human monocytes. *J. Biol. Chem.* **266**, 19490–19498
24. Kim, H. M., Park, B. S., Kim, J.-I., Kim, S. E., Lee, J., Oh, S. C., Enkhbayar, P., Matsushima, N., Lee, H., Yoo, O. J., and Lee, J.-O. (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*. **130**, 906–917
25. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
26. Lewis, L. A., and Ram, S. (2014) Meningococcal disease and the complement system. *Virulence*. **5**, 98–126
27. Gaddy, J. A., Radin, J. N., Cullen, T. W., Chazin, W. J., Skaar, E. P., Trent, M. S., and Algood, H. M. S. (2015) Helicobacter pylori resists the antimicrobial activity of calprotectin via lipid A modification and associated biofilm formation. *mBio*. 10.1128/mBio.01349-15
28. Henderson, J. C., Herrera, C. M., and Trent, M. S. (2017) AlmG, responsible for polymyxin resistance in pandemicVibrio cholerae, is a glycyltransferase distantly related to lipid A late acyltransferases. *J. Biol. Chem.* **292**, 21205–21215
29. Han, M.-L., Velkov, T., Zhu, Y., Roberts, K. D., Le Brun, A. P., Chow, S. H., Gutu, A. D., Moskowitz, S. M., Shen, H.-H., and Li, J. (2018) Polymyxin-induced lipid A

deacylation in Pseudomonas aeruginosa perturbs polymyxin penetration and confers high-level resistance. *ACS Chem. Biol.* **13**, 121–130
30. Bishop, R. E., Gibbons, H. S., Guina, T., Trent, M. S., Miller, S. I., and Raetz, C. R. (2000) Transfer of palmitate from phospholipids to lipid A in outer membranes of gram-negative bacteria. *EMBO J.* **19**, 5071–5080
31. Hittle, L. E., Jones, J. W., Hajjar, A. M., Ernst, R. K., and Preston, A. (2015) Bordetella parapertussis PagP mediates the addition of two palmitates to the lipopolysaccharide lipid A. *J. Bacteriol.* **197**, 572–580
32. Chalabaev, S., Chauhan, A., Novikov, A., Iyer, P., Szczesny, M., Beloin, C., Caroff, M., and Ghigo, J.-M. (2014) Biofilms formed by gram-negative bacteria undergo increased lipid a palmitoylation, enhancing in vivo survival. *mBio*. 10.1128/mBio.01116-14
33. Segev-Zarko, L.-A., Kapach, G., Josten, M., Klug, Y. A., Sahl, H.-G., and Shai, Y. (2018) Deficient lipid A remodeling by the arnB gene promotes biofilm formation in antimicrobial peptide susceptible Pseudomonas aeruginosa. *Biochemistry (Mosc.)*. **57**, 2024–2034
34. Helander Ilkka M., Kilpeläinen Ilkka, and Vaara Martti (2006) Increased substitution of phosphate groups in lipopolysaccharides and lipid A of the polymyxin-resistant pmrA mutants of Salmonella typhimurium: a 31P-NMR study. *Mol. Microbiol.* **11**, 481–487
35. Lee, H., Hsu, F.-F., Turk, J., and Groisman, E. A. (2004) The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and polymyxin resistance in Salmonella enterica. *J. Bacteriol.* **186**, 4124–4133
36. Hankins, J. V., Madsen, J. A., Giles, D. K., Childers, B. M., Klose, K. E., Brodbelt, J. S., and Trent, M. S. (2011) Elucidation of a novel Vibrio cholerae lipid A secondary hydroxy-acyltransferase and its role in innate immune recognition. *Mol. Microbiol.* **81**, 1313–1329
37. Gmeiner, J., Lüderitz, O., and Westphal, O. (1969) Biochemical studies on lipopolysaccharides of Salmonella R mutants. *Eur. J. Biochem.* **7**, 370–379
38. Rosner, M. R., Verret, R. C., and Khorana, H. G. (1979) The structure of lipopolysaccharide from an Escherichia coli heptose-less mutant. III. Two fatty acyl amidases from Dictyostelium discoideum and their action on lipopolysaccharide derivatives. *J. Biol. Chem.* **254**, 5926–5933
39. Strain, S. M., Fesik, S. W., and Armitage, I. M. (1983) Characterization of lipopolysaccharide from a heptoseless mutant of Escherichia coli by carbon 13 nuclear magnetic resonance. *J. Biol. Chem.* **258**, 2906–2910
40. Ray, B. L., Painter, G., and Raetz, C. R. (1984) The biosynthesis of gram-negative endotoxin. Formation of lipid A disaccharides from monosaccharide precursors in extracts of Escherichia coli. *J. Biol. Chem.* **259**, 4852–4859
41. Strain, S. M., Armitage, I. M., Anderson, L., Takayama, K., Qureshi, N., and Raetz, C. R. (1985) Location of polar substituents and fatty acyl chains on lipid A precursors from a 3-deoxy-D-manno-octulosonic acid-deficient mutant of Salmonella typhimurium. Studies by 1H, 13C, and 31P nuclear magnetic resonance. *J. Biol. Chem.* **260**, 16089–16098

42. Gmeiner, J., Simon, M., and Lüderitz, O. (1971) The linkage of phosphate groups and of 2-keto-3-deoxyoctonate to the lipid A component in a Salmonella minnesota lipopolysaccharide. *Eur. J. Biochem.* **21**, 355–356
43. Rosner, M. R., Tang, J., Barzilay, I., and Khorana, H. G. (1979) Structure of the lipopolysaccharide from an Escherichia coli heptose-less mutant. I. Chemical degradations and identification of products. *J. Biol. Chem.* **254**, 5906–5917
44. Wollenweber, H. W., Broady, K. W., Lüderitz, O., and Rietschel, E. T. (1982) The chemical structure of lipid A. Demonstration of amide-linked 3-acyloxyacyl residues in Salmonella minnesota Re lipopolysaccharide. *Eur. J. Biochem.* **124**, 191–198
45. Nishijima, M., and Raetz, C. R. (1979) Membrane lipid biogenesis in Escherichia coli: identification of genetic loci for phosphatidylglycerophosphate synthetase and construction of mutants lacking phosphatidylglycerol. *J. Biol. Chem.* **254**, 7837–7844
46. Nishijima, M., and Raetz, C. R. (1981) Characterization of two membrane-associated glycolipids from an Escherichia coli mutant deficient in phosphatidylglycerol. *J. Biol. Chem.* **256**, 10690–10696
47. Takayama, K., Qureshi, N., Mascagni, P., Nashed, M. A., Anderson, L., and Raetz, C. R. (1983) Fatty acyl derivatives of glucosamine 1-phosphate in Escherichia coli and their relation to lipid A. Complete structure of A diacyl GlcN-1-P found in a phosphatidylglycerol-deficient mutant. *J. Biol. Chem.* **258**, 7379–7385
48. Qureshi, N., Takayama, K., Heller, D., and Fenselau, C. (1983) Position of ester groups in the lipid A backbone of lipopolysaccharides obtained from Salmonella typhimurium. *J. Biol. Chem.* **258**, 12947–12951
49. Seydel, U., Lindner, B., Wollenweber, H. W., and Rietschel, E. T. (1984) Structural studies on the lipid A component of enterobacterial lipopolysaccharides by laser desorption mass spectrometry. Location of acyl groups at the lipid A backbone. *Eur. J. Biochem.* **145**, 505–509
50. Strain, S. M., Fesik, S. W., and Armitage, I. M. (1983) Structure and metal-binding properties of lipopolysaccharides from heptoseless mutants of Escherichia coli studied by 13C and 31P nuclear magnetic resonance. *J. Biol. Chem.* **258**, 13466–13477
51. Raetz, C. R., Purcell, S., Meyer, M. V., Qureshi, N., and Takayama, K. (1985) Isolation and characterization of eight lipid A precursors from a 3-deoxy-D-manno-octylosonic acid-deficient mutant of Salmonella typhimurium. *J. Biol. Chem.* **260**, 16080–16088
52. Brozek, K. A., Hosaka, K., Robertson, A. D., and Raetz, C. R. (1989) Biosynthesis of lipopolysaccharide in Escherichia coli. Cytoplasmic enzymes that attach 3-deoxy-D-manno-octulosonic acid to lipid A. *J. Biol. Chem.* **264**, 6956–6966
53. Belunis, C. J., and Raetz, C. R. (1992) Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from Escherichia coli. *J. Biol. Chem.* **267**, 9988–9997
54. Strain, S. M., and Armitage, I. M. (1985) Selective detection of 3-deoxymannooctulosonic acid in intact lipopolysaccharides by spin-echo 13C NMR. *J. Biol. Chem.* **260**, 12974–12977
55. Zähringer, U., Lindner, B., Seydel, U., Rietschel, E. T., Naoki, H., Unger, F. M., Imoto, M., Kusumoto, S., and Shiba, T. (1985) Structure of de-o-acylated lipopolysaccharide from the escherichia coli re mutant strain F 515. *Tetrahedron Lett.* **26**, 6321–6324

56. Brade, H., and Rietschel, E.-T. (1984) α-2→4-Interlinked 3-deoxy-d-manno-octulosonic acid disaccharide. *Eur. J. Biochem.* **145**, 231–236
57. Anderson, M. S., Bulawa, C. E., and Raetz, C. R. (1985) The biosynthesis of gram-negative endotoxin. Formation of lipid A precursors from UDP-GlcNAc in extracts of Escherichia coli. *J. Biol. Chem.* **260**, 15536–15541
58. Anderson, M. S., and Raetz, C. R. (1987) Biosynthesis of lipid A precursors in Escherichia coli. A cytoplasmic acyltransferase that converts UDP-N-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *J. Biol. Chem.* **262**, 5159–5169
59. Anderson, M. S., Bull, H. G., Galloway, S. M., Kelly, T. M., Mohan, S., Radika, K., and Raetz, C. R. (1993) UDP-N-acetylglucosamine acyltransferase of Escherichia coli. The first step of endotoxin biosynthesis is thermodynamically unfavorable. *J. Biol. Chem.* **268**, 19858–19865
60. Raetz, C. R. H., and Roderick, S. L. (1995) A left-handed parallel β Helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science*. **270**, 997–1000
61. Ulaganathan, V., Buetow, L., and Hunter, W. N. (2007) Nucleotide substrate recognition by UDP-N-acetylglucosamine acyltransferase (LpxA) in the first step of lipid A biosynthesis. *J. Mol. Biol.* **369**, 305–312
62. Williams, A. H., and Raetz, C. R. H. (2007) Structural basis for the acyl chain selectivity and mechanism of UDP-N-acetylglucosamine acyltransferase. *Proc. Natl. Acad. Sci.* **104**, 13543–13550
63. Wyckoff, T. J. O., and Raetz, C. R. H. (1999) The active site of Escherichia coli UDP-N-acetylglucosamine acyltransferase. Chemical modification and site-directed mutagenesis. *J. Biol. Chem.* **274**, 27047–27055
64. Wyckoff, T. J. O., Lin, S., Cotter, R. J., Dotson, G. D., and Raetz, C. R. H. (1998) Hydrocarbon rulers in UDP-N-acetylglucosamine acyltransferases. *J. Biol. Chem.* **273**, 32369–32372
65. Anderson, M. S., Robertson, A. D., Macher, I., and Raetz, C. R. (1988) Biosynthesis of lipid A in Escherichia coli: identification of UDP-3-O-[(R)-3-hydroxymyristoyl]-alpha-D-glucosamine as a precursor of UDP-N2,O3-bis[(R)-3-hydroxymyristoyl]-alpha-D-glucosamine. *Biochemistry (Mosc.)*. **27**, 1908–1917
66. Young, K., Silver, L. L., Bramhill, D., Cameron, P., Eveland, S. S., Raetz, C. R. H., Hyland, S. A., and Anderson, M. S. (1995) The envA permeability/cell division gene of Escherichia coli encodes the second enzyme of lipid A biosynthesis. UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase. *J. Biol. Chem.* **270**, 30384–30391
67. Jackman, J. E., Raetz, C. R. H., and Fierke, C. A. (1999) UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase of Escherichia coli is a zinc metalloenzyme. *Biochemistry (Mosc.)*. **38**, 1902–1911
68. Ogura, T., Inoue, K., Tatsuta, T., Suzaki, T., Karata, K., Young, K., Su, L. H., Fierke, C. A., Jackman, J. E., Raetz, C. R., Coleman, J., Tomoyasu, T., and Matsuzawa, H. (1999) Balanced biosynthesis of major membrane components through regulated degradation of the committed enzyme of lipid A biosynthesis by the AAA protease FtsH (HflB) in Escherichia coli. *Mol. Microbiol.* **31**, 833–844
69. Emiola, A., Andrews, S. S., Heller, C., and George, J. (2016) Crosstalk between the lipopolysaccharide and phospholipid pathways during outer membrane biogenesis in Escherichia coli. *Proc. Natl. Acad. Sci.* **113**, 3108–3113

70. Whittington, D. A., Rusche, K. M., Shin, H., Fierke, C. A., and Christianson, D. W. (2003) Crystal structure of LpxC, a zinc-dependent deacetylase essential for endotoxin biosynthesis. *Proc. Natl. Acad. Sci.* **100**, 8146–8150
71. Clayton, G. M., Klein, D. J., Rickert, K. W., Patel, S. B., Kornienko, M., Zugay-Murphy, J., Reid, J. C., Tummala, S., Sharma, S., Singh, S. B., Miesel, L., Lumb, K. J., and Soisson, S. M. (2013) Structure of the bacterial deacetylase LpxC bound to the nucleotide reaction product reveals mechanisms of oxyanion stabilization and proton transfer. *J. Biol. Chem.* **288**, 34073–34080
72. Coggins, B. E., Li, X., McClerren, A. L., Hindsgaul, O., Raetz, C. R. H., and Zhou, P. (2003) Structure of the LpxC deacetylase with a bound substrate-analog inhibitor. *Nat. Struct. Biol.* **10**, 645–651
73. Jackman, J. E., Raetz, C. R. H., and Fierke, C. A. (2001) Site-directed mutagenesis of the bacterial metalloamidase UDP-(3-O-acyl)-N-acetylglucosamine deacetylase (LpxC). Identification of the zinc binding site. *Biochemistry (Mosc.)*. **40**, 514–523
74. Hernick, M., and Fierke, C. A. (2006) Catalytic mechanism and molecular recognition of E. coli UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase probed by mutagenesis. *Biochemistry (Mosc.)*. **45**, 15240–15248
75. Kelly, T. M., Stachula, S. A., Raetz, C. R., and Anderson, M. S. (1993) The firA gene of Escherichia coli encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase. The third step of endotoxin biosynthesis. *J. Biol. Chem.* **268**, 19866–19874
76. Bartling, C. M., and Raetz, C. R. H. (2008) Steady-state kinetics and mechanism of LpxD, the N-acyltransferase of lipid A biosynthesis. *Biochemistry (Mosc.)*. **47**, 5290–5302
77. Masoudi, A., Raetz, C. R. H., Zhou, P., and Pemble, C. W., 4th (2014) Chasing acyl carrier protein through a catalytic cycle of lipid A production. *Nature*. **505**, 422–426
78. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786
79. Buetow, L., Smith, T. K., Dawson, A., Fyffe, S., and Hunter, W. N. (2007) Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis. *Proc. Natl. Acad. Sci.* **104**, 4321–4326
80. Bartling, C. M., and Raetz, C. R. H. (2009) Crystal structure and acyl chain selectivity of Escherichia coli LpxD, the N-acyltransferase of lipid A biosynthesis. *Biochemistry (Mosc.)*. **48**, 8672–8683
81. Babinski, K. J., Ribeiro, A. A., and Raetz, C. R. H. (2002) The Escherichia coli gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *J. Biol. Chem.* **277**, 25937–25946
82. Metzger, L. E., 4th, and Raetz, C. R. H. (2010) An alternative route for UDP-diacylglucosamine hydrolysis in bacterial lipid A biosynthesis. *Biochemistry (Mosc.)*. **49**, 6715–6726
83. Young, H. E., Zhao, J., Barker, J. R., Guan, Z., Valdivia, R. H., and Zhou, P. (2016) Discovery of the elusive UDP-diacylglucosamine hydrolase in the lipid A biosynthetic pathway in Chlamydia trachomatis. *mBio*. 10.1128/mBio.00090-16

84. Young, H. E., Donohue, M. P., Smirnova, T. I., Smirnov, A. I., and Zhou, P. (2013) The UDP-diacylglucosamine pyrophosphohydrolase LpxH in lipid A biosynthesis utilizes Mn2+ cluster for catalysis. *J. Biol. Chem.* **288**, 26987–27001
85. Okada, C., Wakabayashi, H., Kobayashi, M., Shinoda, A., Tanaka, I., and Yao, M. (2016) Crystal structures of the UDP-diacylglucosamine pyrophosphohydrase LpxH from Pseudomonas aeruginosa. *Sci. Rep.* 10.1038/srep32822
86. Babinski, K. J., Kanjilal, S. J., and Raetz, C. R. H. (2002) Accumulation of the lipid A precursor UDP-2,3-diacylglucosamine in an Escherichia coli mutant lacking the lpxH gene. *J. Biol. Chem.* **277**, 25947–25956
87. Metzger, L. E., 4th, Lee, J. K., Finer-Moore, J. S., Raetz, C. R. H., and Stroud, R. M. (2012) LpxI structures reveal how a lipid A precursor is synthesized. *Nat. Struct. Mol. Biol.* **19**, 1132–1138
88. Ray, B. L., and Raetz, C. R. (1987) The biosynthesis of gram-negative endotoxin. A novel kinase in Escherichia coli membranes that incorporates the 4'-phosphate of lipid A. *J. Biol. Chem.* **262**, 1122–1128
89. Garrett, T. A., Kadrmas, J. L., and Raetz, C. R. H. (1997) Identification of the gene encoding the Escherichia coli lipid A 4'-kinase. Facile phosphorylation of endotoxin analogs with recombinant LpxK. *J. Biol. Chem.* **272**, 21855–21864
90. Emptage, R. P., Pemble, C. W., York, J. D., Raetz, C. R. H., and Zhou, P. (2013) Mechanistic characterization of the tetraacyldisaccharide-1-phosphate 4'-kinase LpxK involved in lipid A biosynthesis. *Biochemistry (Mosc.)*. **52**, 2280–2290
91. Emptage, R. P., Daughtry, K. D., Pemble, C. W., and Raetz, C. R. H. (2012) Crystal structure of LpxK, the 4'-kinase of lipid A biosynthesis and atypical P-loop kinase functioning at the membrane interface. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12956–12961
92. Emptage, R. P., Tonthat, N. K., York, J. D., Schumacher, M. A., and Zhou, P. (2014) Structural basis of lipid binding for the membrane-embedded tetraacyldisaccharide-1-phosphate 4'-kinase LpxK. *J. Biol. Chem.* **289**, 24059–24068
93. Brabetz, W., Lindner, B., and Brade, H. (2000) Comparative analyses of secondary gene products of 3-deoxy-D-manno-oct-2-ulosonic acid transferases from Chlamydiaceae in Escherichia coli K-12. *Eur. J. Biochem.* **267**, 5458–5465
94. Mamat, U., Schmidt, H., Munoz, E., Lindner, B., Fukase, K., Hanuszkiewicz, A., Wu, J., Meredith, T. C., Woodard, R. W., Hilgenfeld, R., Mesters, J. R., and Holst, O. (2009) WaaA of the hyperthermophilic bacterium Aquifex aeolicus is a monofunctional 3-deoxy-D-manno-oct-2-ulosonic acid transferase involved in lipopolysaccharide biosynthesis. *J. Biol. Chem.* **284**, 22248–22262
95. Katz, C., and Ron, E. Z. (2008) Dual role of FtsH in regulating lipopolysaccharide biosynthesis in Escherichia coli. *J. Bacteriol.* **190**, 7117–7122
96. Schmidt, H., Hansen, G., Singh, S., Hanuszkiewicz, A., Lindner, B., Fukase, K., Woodard, R. W., Holst, O., Hilgenfeld, R., Mamat, U., and Mesters, J. R. (2012) Structural and mechanistic analysis of the membrane-embedded glycosyltransferase WaaA required for lipopolysaccharide synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6253–6258
97. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) The 1.9 A crystal structure of Escherichia coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Sci.* **9**, 1045–1052

98. Albesa-Jové, D., Giganti, D., Jackson, M., Alzari, P. M., and Guerin, M. E. (2014) Structure–function relationships of membrane-associated GT-B glycosyltransferases. *Glycobiology*. **24**, 108–124
99. Guerin, M. E., Kordulakova, J., Schaeffer, F., Svetlikova, Z., Buschiazzo, A., Giganti, D., Gicquel, B., Mikusova, K., Jackson, M., and Alzari, P. M. (2007) Molecular recognition and interfacial catalysis by the essential phosphatidylinositol mannosyltransferase PimA from mycobacteria. *J. Biol. Chem.* **282**, 20705–20714
100. Rodrigo-Unzueta, A., Martínez, M. A., Comino, N., Alzari, P. M., Chenal, A., and Guerin, M. E. (2016) Molecular basis of membrane association by the phosphatidylinositol mannosyltransferase PimA enzyme from mycobacteria. *J. Biol. Chem.* **291**, 13955–13963
101. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 845–849
102. Giganti, D., Albesa-Jové, D., Urresti, S., Rodrigo-Unzueta, A., Martínez, M. A., Comino, N., Barilone, N., Bellinzoni, M., Chenal, A., Guerin, M. E., and Alzari, P. M. (2015) Secondary structure reshuffling modulates glycosyltransferase function at the membrane. *Nat. Chem. Biol.* **11**, 16–18
103. Clementz, T., Bednarski, J. J., and Raetz, C. R. H. (1996) Function of the htrB high temperature requirement gene of Escherichia coli in the acylation of lipid A: HtrB catalyzed incorporation of laurate. *J. Biol. Chem.* **271**, 12095–12102
104. Six, D. A., Carty, S. M., Guan, Z., and Raetz, C. R. H. (2008) Purification and mutagenesis of LpxL, the lauroyltransferase of Escherichia coli lipid A biosynthesis. *Biochemistry (Mosc.)*. **47**, 8623–8637
105. Carty, S. M., Sreekumar, K. R., and Raetz, C. R. H. (1999) Effect of cold shock on lipid A biosynthesis in Escherichia coli. Induction At 12 °C of an acyltransferase specific for palmitoleoyl-acyl carrier protein. *J. Biol. Chem.* **274**, 9677–9685
106. Vorachek-Warren, M. K., Carty, S. M., Lin, S., Cotter, R. J., and Raetz, C. R. H. (2002) An Escherichia coli mutant lacking the cold shock-induced palmitoleoyltransferase of lipid A biosynthesis: absence of unsaturated acyl chains and antibiotic hypersensitivity at 12 °C. *J. Biol. Chem.* **277**, 14186–14193
107. Dovala, D., Rath, C. M., Hu, Q., Sawyer, W. S., Shia, S., Elling, R. A., Knapp, M. S., and Metzger, L. E. (2016) Structure-guided enzymology of the lipid A acyltransferase LpxM reveals a dual activity mechanism. *Proc. Natl. Acad. Sci.* **113**, E6064–E6071
108. Clementz, T., Zhou, Z., and Raetz, C. R. H. (1997) Function of the Escherichia coli msbB gene, a multicopy suppressor of htrB knockouts, in the acylation of lipid A. Acylation by MsbB follows laurate incorporation by HtrB. *J. Biol. Chem.* **272**, 10353–10360
109. Röttig, A., and Steinbüchel, A. (2013) Acyltransferases in bacteria. *Microbiol. Mol. Biol. Rev.* **77**, 277–321
110. Tamada, T., Feese, M. D., Ferri, S. R., Kato, Y., Yajima, R., Toguri, T., and Kuroki, R. (2004) Substrate recognition and selectivity of plant glycerol-3-phosphate acyltransferases (GPATs) from Cucurbita moscata and Spinacea oleracea. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 13–21

111. Rubin, E. J., O'Brien, J. P., Ivanov, P. L., Brodbelt, J. S., and Trent, M. S. (2014) Identification of a broad family of lipid A late acyltransferases with non-canonical substrate specificity. *Mol. Microbiol.* **91**, 887–899
112. Whitfield, C., and Trent, M. S. (2014) Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* **83**, 99–128
113. Mudapaka, J., and Taylor, E. A. (2015) Cloning and characterization of the Escherichia coli Heptosyltransferase III: Exploring substrate specificity in lipopolysaccharide core biosynthesis. *FEBS Lett.* **589**, 1423–1429
114. Yethon, J. A., Vinogradov, E., Perry, M. B., and Whitfield, C. (2000) Mutation of the lipopolysaccharide core glycosyltransferase encoded by waaG destabilizes the outer membrane of Escherichia coli by interfering with core phosphorylation. *J. Bacteriol.* **182**, 5620–5623
115. Qian, J., Garrett, T. A., and Raetz, C. R. H. (2014) In vitro assembly of the outer core of the lipopolysaccharide from Escherichia coli K-12 and Salmonella typhimurium. *Biochemistry (Mosc.)*. **53**, 1250–1262
116. Reynolds, C. M., Kalb, S. R., Cotter, R. J., and Raetz, C. R. H. (2005) A phosphoethanolamine transferase specific for the outer 3-deoxy-D-manno-octulosonic acid residue of Escherichia coli lipopolysaccharide. Identification of the eptB gene and Ca2+ hypersensitivity of an eptB deletion mutant. *J. Biol. Chem.* **280**, 21202–21211
117. Heinrichs, D. E., Yethon, J. A., and Whitfield, C. (1998) Molecular basis for structural diversity in the core regions of the lipopolysaccharides of Escherichia coli and Salmonella enterica. *Mol. Microbiol.* **30**, 221–232
118. Meredith, T. C., Mamat, U., Kaczynski, Z., Lindner, B., Holst, O., and Woodard, R. W. (2007) Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in Escherichia coli. *J. Biol. Chem.* **282**, 7790–7798
119. Doerrler, W. T., Gibbons, H. S., and Raetz, C. R. H. (2004) MsbA-dependent translocation of lipids across the inner membrane of Escherichia coli. *J. Biol. Chem.* **279**, 45102–45109
120. Ward, A., Reyes, C. L., Yu, J., Roth, C. B., and Chang, G. (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc. Natl. Acad. Sci.* **104**, 19005–19010
121. Mi, W., Li, Y., Yoon, S. H., Ernst, R. K., Walz, T., and Liao, M. (2017) Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature*. **549**, 233–237
122. Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu. Rev. Biochem.* **75**, 39–68
123. Rick, P. D., Barr, K., Sankaran, K., Kajimura, J., Rush, J. S., and Waechter, C. J. (2003) Evidence that the wzxE gene of Escherichia coli K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen. *J. Biol. Chem.* **278**, 16534–16542
124. Islam, S. T., Taylor, V. L., Qi, M., and Lam, J. S. (2010) Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from Pseudomonas aeruginosa PAO1 reveals novel domain architectures. *mBio*. 10.1128/mBio.00189-10
125. Ruan, X., Loyola, D. E., Marolda, C. L., Perez-Donoso, J. M., and Valvano, M. A. (2012) The WaaL O-antigen lipopolysaccharide ligase has features in common with metal ion-independent inverting glycosyltransferases. *Glycobiology*. **22**, 288–299

126. Hong, Y., Liu, M. A., and Reeves, P. R. (2018) Progress in our understanding of Wzx flippase for translocation of bacterial membrane lipid-linked oligosaccharide. *J. Bacteriol.* 10.1128/JB.00154-17
127. Islam, S. T., Gold, A. C., Taylor, V. L., Anderson, E. M., Ford, R. C., and Lam, J. S. (2011) Dual conserved periplasmic loops possess essential charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in Pseudomonas aeruginosa PAO1. *J. Biol. Chem.* **286**, 20600–20605
128. Islam, S. T., Huszczynski, S. M., Nugent, T., Gold, A. C., and Lam, J. S. (2013) Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1. *Sci. Rep.* 10.1038/srep03441
129. Kalynych, S., Yao, D., Magee, J., and Cygler, M. (2012) Structural characterization of closely related O-antigen lipopolysaccharide (LPS) chain length regulators. *J. Biol. Chem.* **287**, 15696–15705
130. Hagelueken, G., Clarke, B. R., Huang, H., Tuukkanen, A., Danciu, I., Svergun, D. I., Hussain, R., Liu, H., Whitfield, C., and Naismith, J. H. (2015) A coiled-coil domain acts as a molecular ruler to regulate O-antigen chain length in lipopolysaccharide. *Nat. Struct. Mol. Biol.* **22**, 50–56
131. Clarke, B. R., Greenfield, L. K., Bouwman, C., and Whitfield, C. (2009) Coordination of polymerization, chain termination, and export in assembly of the Escherichia coli lipopolysaccharide O9a antigen in an ATP-binding cassette transporter-dependent pathway. *J. Biol. Chem.* **284**, 30662–30672
132. Tocilj, A., Munger, C., Proteau, A., Morona, R., Purins, L., Ajamian, E., Wagner, J., Papadopoulos, M., Van Den Bosch, L., Rubinstein, J. L., Féthière, J., Matte, A., and Cygler, M. (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat. Struct. Mol. Biol.* **15**, 130–138
133. Collins, R. F., Kargas, V., Clarke, B. R., Siebert, C. A., Clare, D. K., Bond, P. J., Whitfield, C., and Ford, R. C. (2017) Full-length, oligomeric structure of Wzz determined by cryoelectron microscopy reveals insights into membrane-bound states. *Structure*. **25**, 806–815
134. Greenfield, L. K., and Whitfield, C. (2012) Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways. *Carbohydr. Res.* **356**, 12–24
135. Cuthbertson, L., Kimber, M. S., and Whitfield, C. (2007) Substrate binding by a bacterial ABC transporter involved in polysaccharide export. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19529–19534
136. Bi, Y., Mann, E., Whitfield, C., and Zimmer, J. (2018) Architecture of a channel-forming O-antigen polysaccharide ABC transporter. *Nature*. **553**, 361–365
137. Ma, B., Reynolds, C. M., and Raetz, C. R. H. (2008) Periplasmic orientation of nascent lipid A in the inner membrane of an Escherichia coli LptA mutant. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13823–13828
138. Chng, S.-S., Gronenberg, L. S., and Kahne, D. (2010) Proteins required for lipopolysaccharide assembly in Escherichia coli form a transenvelope complex. *Biochemistry (Mosc.)*. **49**, 4565–4567
139. Okuda, S., Freinkman, E., and Kahne, D. (2012) Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. *Science*. **338**, 1214–1217

140. Sherman, D. J., Xie, R., Taylor, R. J., George, A. H., Okuda, S., Foster, P. J., Needleman, D. J., and Kahne, D. (2018) Lipopolysaccharide is transported to the cell surface by a membrane-to-membrane protein bridge. *Science*. **359**, 798–801
141. Villa, R., Martorana, A. M., Okuda, S., Gourlay, L. J., Nardini, M., Sperandeo, P., Dehò, G., Bolognesi, M., Kahne, D., and Polissi, A. (2013) The Escherichia coli Lpt transenvelope protein complex for lipopolysaccharide export is assembled via conserved structurally homologous domains. *J. Bacteriol.* **195**, 1100–1108
142. Sperandeo, P., Villa, R., Martorana, A. M., Samalikova, M., Grandori, R., Dehò, G., and Polissi, A. (2011) New insights into the Lpt machinery for lipopolysaccharide transport to the cell surface: LptA-LptC interaction and LptA stability as sensors of a properly assembled transenvelope complex. *J. Bacteriol.* **193**, 1042–1053
143. Freinkman, E., Okuda, S., Ruiz, N., and Kahne, D. (2012) Regulated assembly of the transenvelope protein complex required for lipopolysaccharide export. *Biochemistry (Mosc.)*. **51**, 4800–4806
144. Luo, Q., Yang, X., Yu, S., Shi, H., Wang, K., Xiao, L., Zhu, G., Sun, C., Li, T., Li, D., Zhang, X., Zhou, M., and Huang, Y. (2017) Structural basis for lipopolysaccharide extraction by ABC transporter LptB2FG. *Nat. Struct. Mol. Biol.* **24**, 469–474
145. Suits, M. D. L., Sperandeo, P., Dehò, G., Polissi, A., and Jia, Z. (2008) Novel structure of the conserved gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *J. Mol. Biol.* **380**, 476–488
146. Tran, A. X., Dong, C., and Whitfield, C. (2010) Structure and functional analysis of LptC, a conserved membrane protein involved in the lipopolysaccharide export pathway in Escherichia coli. *J. Biol. Chem.* **285**, 33529–33539
147. Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C., and Huang, Y. (2014) Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature*. **511**, 108–111
148. Dong, H., Zhang, Z., Tang, X., Paterson, N. G., and Dong, C. (2017) Structural and functional insights into the lipopolysaccharide ABC transporter LptB2FG. *Nat. Commun.* **8**, 222
149. Tefsen, B., Geurtsen, J., Beckers, F., Tommassen, J., and de Cock, H. (2005) Lipopolysaccharide transport to the bacterial outer membrane in spheroplasts. *J. Biol. Chem.* **280**, 4504–4509
150. Sherman, D. J., Lazarus, M. B., Murphy, L., Liu, C., Walker, S., Ruiz, N., and Kahne, D. (2014) Decoupling catalytic activity from biological function of the ATPase that powers lipopolysaccharide transport. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4982–4987
151. Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W., and Dong, C. (2014) Structural basis for outer membrane lipopolysaccharide insertion. *Nature*. **511**, 52–56
152. Malojčić, G., Andres, D., Grabowicz, M., George, A. H., Ruiz, N., Silhavy, T. J., and Kahne, D. (2014) LptE binds to and alters the physical state of LPS to catalyze its assembly at the cell surface. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9467–9472
153. Bulawa, C. E., and Raetz, C. R. (1984) The biosynthesis of gram-negative endotoxin. Identification and function of UDP-2,3-diacylglucosamine in Escherichia coli. *J. Biol. Chem.* **259**, 4846–4851

154. Cho, J., Lee, C.-J., Zhao, J., Young, H. E., and Zhou, P. (2016) Structure of the essential Haemophilus influenzae UDP-diacylglucosamine pyrophosphohydrolase LpxH in lipid A biosynthesis. *Nat. Microbiol.* 10.1038/nmicrobiol.2016.154
155. Mitić, N., Smith, S. J., Neves, A., Guddat, L. W., Gahan, L. R., and Schenk, G. (2006) The catalytic mechanisms of binuclear metallohydrolases. *Chem. Rev.* **106**, 3338–3363
156. Nayar, A. S., Dougherty, T. J., Ferguson, K. E., Granger, B. A., McWilliams, L., Stacey, C., Leach, L. J., Narita, S.-I., Tokuda, H., Miller, A. A., Brown, D. G., and McLeod, S. M. (2015) Novel antibacterial targets and compounds revealed by a high-throughput cell wall reporter assay. *J. Bacteriol.* **197**, 1726–1734
157. Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461
158. Felli, I. C., and Brutscher, B. (2009) Recent advances in solution NMR: fast methods and heteronuclear direct detection. *ChemPhysChem*. **10**, 1356–1368
159. Klare, J. P., and Steinhoff, H.-J. (2009) Spin labeling EPR. *Photosynth. Res.* **102**, 377–390
160. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858
161. Podobnik, M., Tyagi, R., Matange, N., Dermol, U., Gupta, A. K., Mattoo, R., Seshadri, K., and Visweswariah, S. S. (2009) A mycobacterial cyclic AMP phosphodiesterase that moonlights as a modifier of cell wall permeability. *J. Biol. Chem.* **284**, 32846–32857
162. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S., and Eisenberg, D. (2007) Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci. Publ. Protein Soc.* **16**, 1569–1576
163. Doublié, S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Macromol. Crystallogr. Protoc.* **363**, 91–108
164. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
165. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
166. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
167. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
168. Sours, K. M., and Ahn, N. G. (2010) Analysis of MAP kinases by hydrogen exchange mass spectrometry. *MAP Kinase Signal. Protoc.* **661**, 239–255
169. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida,

M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A. J., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2016) *Gaussian 09*, Gaussian, Inc., Wallingford CT

170. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, and P.A. Kollman (2014) *AMBER 14*, University of California, San Francisco

171. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174

172. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713

173. Mark, P., and Nilsson, L. (2001) Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A*. **105**, 9954–9960

174. Sondergaard, C. R., Olsson, M. H., Rostkowski, M., and Jensen, J. H. (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J Chem Theory Comput*. **7**, 2284–95

175. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*. **32**, W665-7

176. Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., and Walker, R. C. (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J Chem Theory Comput*. **9**, 3878–88

177. Roe, D. R., and Cheatham, T. E., 3rd (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput*. **9**, 3084–95

178. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph*. **14**, 33–38

179. R Core Team (2018) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria

180. Rodríguez, S., Cesio, M. V., Heinzen, H., and Moyna, P. (2000) Determination of the phospholipid/lipophilic compounds ratio in liposomes by thin-layer chromatography scanning densitometry. *Lipids*. **35**, 1033–1036

181. Schariter, J. A., Pachuski, J., Fried, B., and Sherma, J. (2002) Determination of neutral lipids and phospholipids in the cercariae of Schistosoma mansoni by high

performance thin layer chromatography. *J. Liq. Chromatogr. Relat. Technol.* **25**, 1615–1622
182. Sanner, M. F. (1999) Python: a programming language for software integration and development. *J. Mol. Graph. Model.* **17**, 57–61
183. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791
184. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729
185. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675
186. Metzger, L. E., 4th, and Raetz, C. R. H. (2009) Purification and characterization of the lipid A disaccharide synthase (LpxB) from Escherichia coli, a peripheral membrane protein. *Biochemistry (Mosc.)*. **48**, 11559–11571
187. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555
188. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 10.1038/msb4100050
189. Nishijima, M., Bulawa, C. E., and Raetz, C. R. (1981) Two interacting mutations causing temperature-sensitive phosphatidylglycerol synthesis in Escherichia coli membranes. *J. Bacteriol.* **145**, 113–121
190. Liang, X., Lee, C.-J., Zhao, J., Toone, E. J., and Zhou, P. (2013) Synthesis, structure, and antibiotic activity of aryl-substituted LpxC inhibitors. *J. Med. Chem.* **56**, 6954–6966
191. Liang, X., Lee, C.-J., Chen, X., Chung, H. S., Zeng, D., Raetz, C. R. H., Li, Y., Zhou, P., and Toone, E. J. (2011) Syntheses, structures and antibiotic activities of LpxC inhibitors based on the diacetylene scaffold. *Bioorg. Med. Chem.* **19**, 852–860
192. Löppenberg, M., Müller, H., Pulina, C., Oddo, A., Teese, M., Jose, J., and Holl, R. (2013) Synthesis and biological evaluation of flexible and conformationally constrained LpxC inhibitors. *Org. Biomol. Chem.* **11**, 6056–6070
193. Mansoor, U. F., Vitharana, D., Reddy, P. A., Daubaras, D. L., McNicholas, P., Orth, P., Black, T., and Siddiqui, M. A. (2011) Design and synthesis of potent Gram-negative specific LpxC inhibitors. *Bioorg. Med. Chem. Lett.* **21**, 1155–1161
194. McAllister, L. A., Montgomery, J. I., Abramite, J. A., Reilly, U., Brown, M. F., Chen, J. M., Barham, R. A., Che, Y., Chung, S. W., Menard, C. A., Mitton-Fry, M., Mullins, L. M., Noe, M. C., O'Donnell, J. P., Oliver, R. M., 3rd, Penzien, J. B., Plummer, M., Price, L. M., Shanmugasundaram, V., Tomaras, A. P., and Uccello, D. P. (2012) Heterocyclic methylsulfone hydroxamic acid LpxC inhibitors as Gram-negative antibacterial agents. *Bioorg. Med. Chem. Lett.* **22**, 6832–6838
195. Montgomery, J. I., Brown, M. F., Reilly, U., Price, L. M., Abramite, J. A., Arcari, J., Barham, R., Che, Y., Chen, J. M., Chung, S. W., Collantes, E. M., Desbonnet, C.,

Doroski, M., Doty, J., Engtrakul, J. J., Harris, T. M., Huband, M., Knafels, J. D., Leach, K. L., Liu, S., Marfat, A., McAllister, L., McElroy, E., Menard, C. A., Mitton-Fry, M., Mullins, L., Noe, M. C., O'Donnell, J., Oliver, R., Penzien, J., Plummer, M., Shanmugasundaram, V., Thoma, C., Tomaras, A. P., Uccello, D. P., Vaz, A., and Wishka, D. G. (2012) Pyridone methylsulfone hydroxamate LpxC inhibitors for the treatment of serious gram-negative infections. *J. Med. Chem.* **55**, 1662–1670
196. Warmus, J. S., Quinn, C. L., Taylor, C., Murphy, S. T., Johnson, T. A., Limberakis, C., Ortwine, D., Bronstein, J., Pagano, P., Knafels, J. D., Lightle, S., Mochalkin, I., Brideau, R., and Podoll, T. (2012) Structure based design of an in vivo active hydroxamic acid inhibitor of P. aeruginosa LpxC. *Bioorg. Med. Chem. Lett.* **22**, 2536–2543
197. Jenkins, R. J., and Dotson, G. D. (2012) Dual targeting antibacterial peptide inhibitor of early lipid A biosynthesis. *ACS Chem. Biol.* **7**, 1170–1177
198. Williams, A. H., Immormino, R. M., Gewirth, D. T., and Raetz, C. R. H. (2006) Structure of UDP-N-acetylglucosamine acyltransferase with a bound antibacterial pentadecapeptide. *Proc. Natl. Acad. Sci.* **103**, 10877–10882
199. Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C., and Stuckey, P. J. (2011) Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics*. **27**, 3315–3316
200. Vetting, M. W., Frantom, P. A., and Blanchard, J. S. (2008) Structural and enzymatic analysis of MshA from Corynebacterium glutamicum: substrate-assisted catalysis. *J. Biol. Chem.* **283**, 15834–15844
201. Sobhanifar, S., Worrall, L. J., Gruninger, R. J., Wasney, G. A., Blaukopf, M., Baumann, L., Lameignere, E., Solomonson, M., Brown, E. D., Withers, S. G., and Strynadka, N. C. J. (2015) Structure and mechanism of Staphylococcus aureus TarM, the wall teichoic acid α-glycosyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E576-585
202. Lin, L. Y.-C., Rakic, B., Chiu, C. P. C., Lameignere, E., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. J. (2011) Structure and mechanism of the lipooligosaccharide sialyltransferase from Neisseria meningitidis. *J. Biol. Chem.* **286**, 37237–37248
203. Okuda, S., and Yoshizawa, A. C. (2011) ODB: a database for operon organizations, 2011 update. *Nucleic Acids Res.* **39**, D552-555
204. Ni, L., Chokhawala, H. A., Cao, H., Henning, R., Ng, L., Huang, S., Yu, H., Chen, X., and Fisher, A. J. (2007) Crystal structures of Pasteurella multocida sialyltransferase complexes with acceptor and donor analogues reveal substrate binding sites and catalytic mechanism. *Biochemistry (Mosc.)*. **46**, 6288–6298
205. Vyplel, H., Scholz, D., Macher, I., Schindlmaier, K., and Schuetze, E. (1991) C-Glycosidic analogs of lipid A and lipid X: synthesis and biological activities. *J. Med. Chem.* **34**, 2759–2767
206. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
207. Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of Escherichia coli with plasmids. *Gene*. **96**, 23–28